ISSN 0256-1840

INTERNATIONAL SOCIETY FOR PHOTOGRAMMETRY AND REMOTE SENSING
SOCIÉTÉ INTERNATIONALE DE PHOTOGRAMMÉTRIE ET DE TÉLÉDÉTECTION
INTERNATIONALE GESELLSCHAFT FÜR PHOTOGRAMMETRIE UND FERNERKUNDUNG

INTERNATIONAL ARCHIVES OF PHOTOGRAMMETRY AND REMOTE SENSING
ATIONALES DE PHOTOGRAMMÉTRIE ET DE TÉLÉDÉTECTION
ARCHIV FÜR PHOTOGRAMMETRIE UND FERNERKUNDUNG

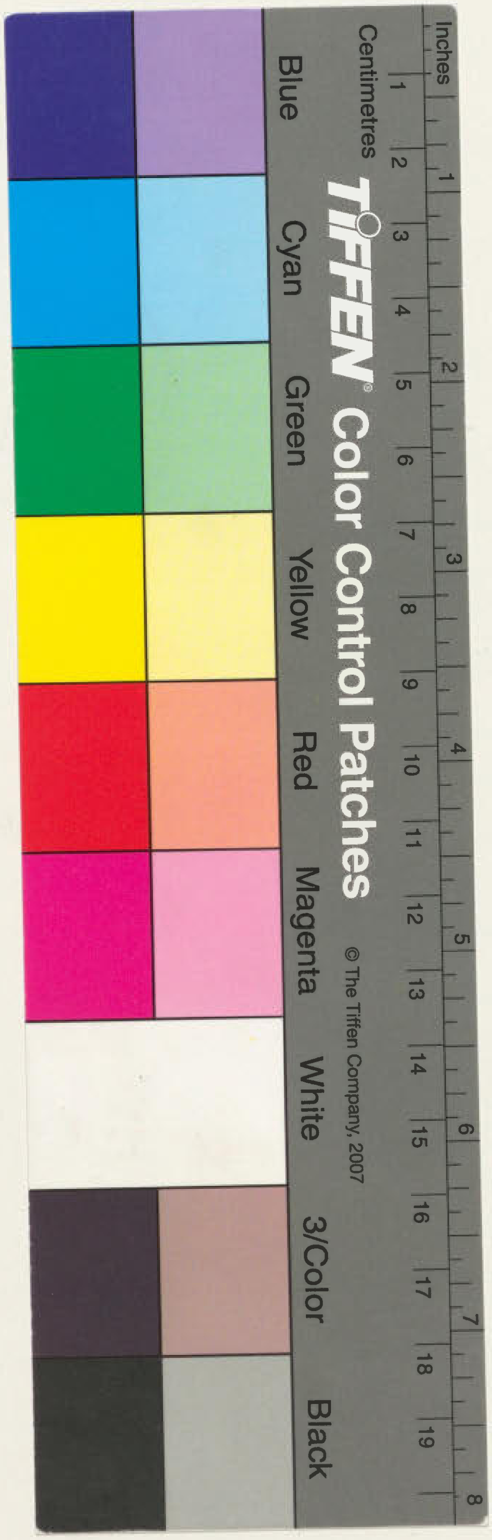
Volume / Volume / Band **30** Part / Tome / Teil **2**

ISPRS Commission II Symposium

SYSTEMS FOR DATA PROCESSING, ANALYSIS AND REPRESENTATION

June 6 -10, 1994
Ottawa, Canada

Published by: The Surveys, Mapping and Remote Sensing Sector, Natural Resources Canada
Publié par : le Secteur des levés, de la cartographie et de la télédétection, Ressources naturelles Canada


Inches
Centimetres
TIFFEN® Color Control Patches
Blue
Cyan
Green
Yellow
Red
Magenta
White
3/Color
Black
© The Tiffen Company, 2007

INTERNATIONAL SOCIETY FOR PHOTOGRAMMETRY AND REMOTE SENSING
SOCIÉTÉ INTERNATIONALE DE PHOTOGRAMMÉTRIE ET DE TÉLÉDÉTECTION
INTERNATIONALE GESELLSCHAFT FÜR PHOTOGRAMMETRIE UND FERNERKUNDUNG

INTERNATIONAL ARCHIVES OF PHOTOGRAMMETRY AND REMOTE SENSING
ARCHIVES INTERNATIONALES DE PHOTOGRAMMÉTRIE ET DE TÉLÉDÉTECTION
INTERNATIONALES ARCHIV FÜR PHOTOGRAMMETRIE UND FERNERKUNDUNG

Volume / Volume / Band **30** Part / Tome / Teil **2**

ISPRS Commission II Symposium

SYSTEMS FOR DATA PROCESSING, ANALYSIS AND REPRESENTATION

June 6 -10, 1994
Ottawa, Canada

TOC

Edited by: Dr. Mosaad Allam, ISPRS Commission II President
Gordon Plunkett, ISPRS Commission II Secretary

Published by: The Surveys, Mapping and Remote Sensing Sector, Natural Resources Canada
Publié par : le Secteur des levés, de la cartographie et de la télédétection, Ressources naturelles Canada

Preface

Early during the planning stages for the Symposium of ISPRS Commission II, it became readily apparent that our Call for Papers was attracting excellent abstract submissions. With these in mind, and with the assistance of the Commission II WG Chairpersons, we eagerly began scheduling plenary, concurrent and poster presentations into a finalized Technical Program in anticipation of June 1994. In completing our responsibilities of hosting the Commission II Symposium, we are pleased to publish these Proceedings as a lasting record of the Commission II Symposium.

As in the case when any large meeting and symposium is assembled, there are many people who work to bring everything together, and a number of these merit special mention. In building this Symposium, we must begin by acknowledging the support of the many authors who submitted papers and made presentations. Without their efforts we would not have had such a successful meeting, and these proceedings are a collection of their efforts in an unedited and original form.

Next, we must salute the efforts of the Technical Program Committee, who had the task of reviewing, organizing and commenting on the presentations for publication. A list of the Technical Program Committee is included in the next few pages and should be accorded your interest.

Our great appreciation is also extended to the ISPRS Symposium Organizing Committee, who more than a year ago began preparations for this Symposium. Their efforts are acknowledged in the entire production, ranging from advance publicity, registration and paper submission to site organization, activity planning and publishing these Proceedings.

As a participant at the ISPRS Commission II Symposium, we hope you benefit from your visit through greater knowledge, friendship and experiences. We wish you a most pleasant stay in Ottawa.

Dr. Mosaad Allam
President
ISPRS Commission II

Gordon Plunkett
Secretary
ISPRS Commission II

ISPRS TECHNICAL COMMITTEE

Gordon Plunkett
Technical Program Director,
SMRSS/NRCan

Mosaad Allam
SMRSS/NRCan

J.A.R. Blais
The University of Calgary

Michael Butler
Land Registration and Information Service

Ian Dowman
University College London
U.K.

Manfred Ehlers
University of Osnabruck-Vechta
Germany

Atef Elassal
National Oceanic and Atmospheric Administration
U.S.A.

Terry Fisher
SMRSS/NRCan

Jeff Labonte
SMRSS/NRCan

Kurt Novak
Ohio State University
U.S.A.

Robert O'Neil
SMRSS/NRCan

Ekow Otoo
Carleton University

Julia Poore
Ph. D. Associates Inc.

Klaus Szangolies
ISPRS Commission II
Germany

COMITÉ TECHNIQUE ISPRS

Gordon Plunkett
Directeur du programme technique
SLCT/RNCan

Mosaad Allam
SLCT/RNCan

J. A. R. Blais
Université de Calgary

Michael Butler
Service du cadastre et de l'information foncière

Ian Dowman
Collège universitaire de Londres
Royaume-Uni

Manfred Ehlers
Université d'Osnabrück-Vechta
Allemagne

Atef Elassal
National Oceanic and Atmospheric Administration
É.-U.

Terry Fisher
SLCT/RNCan

Jeff Labonte
SLCT/RNCan

Kurt Novak
Ohio State University
É.-U.

Robert O'Neil
SLCT/RNCan

Ekow Otoo
Université Carleton

Julia Poore
Ph. D. Associates Inc.

Klaus Szangolies
Commission II de l'ISPRS
Allemagne

Commission II Terms of Reference and Working Groups

President

Mosaad Allam
Tel. + 1 (613) 996-2810
Fax + 1 (613) 952-0916
E-mail Mosaad.Allam@X400.emr.ca

Address:
Geographic Information Systems Division
Surveys, Mapping and Remote Sensing Sector
615 Booth Street, Room 753
Ottawa, Ontario, CANADA
K1A 0E9

Secretary

Gordon Plunkett
Tel. + 1 (613) 992-0389
Fax + 1 (613) 952-0916

Address: Same

Terms of Reference

° Design and development of integrated for measurement, processing, analysis, representation and storage of photogrammetric, remote sensing, and GIS data

° Study and evaluation of system integration aspects for photogrammetry, remote sensing and GIS data processing

° Analysis of systems and their components for automated, semiautomated and manual digital processing systems

° Development of systems and technologies for radar data processing

° Study of real time mapping technologies

° Standardization of digital systems for photogrammetry, remote sensing and GIS

Working Groups

WG II/1 - Real Time Mapping Technologies

Chairperson	Dr. Kurt Novak	U.S.A.
Co-chairperson	Michael Hahn	Germany
Secretary	Holger Schade	Germany
Sessions:	C-1, F-1, G-1, L-1	

WG II/2 - Hardware and Software Aspects of GIS

Chairperson	Dr. Manfred Ehlers	Germany
Co-chairperson	Nickolas Faust	U.S.A.
Secretary	David Steiner	Germany
Sessions:	D-2, J-1	

WG II/3 - Technologies for Large Volumes of Spatial Data

Chairperson	Dr. Ekow Otoo	Canada
Co-chairperson	Terry Fisher	Canada
Secretary	Cherian Chaly	Canada
Sessions:	B-1, F-2, I-1	

WG II/4 - Systems for the Processing of Radar Data

Chairperson	Dr. Robert O'Neil	Canada
Co-chairperson	Dr. Hiroshi Kimura	Japan
Secretary	Dr. Marc D'Iorio	Canada
Sessions:	A-1, K-1	

WG II/5 - Integrated Production Systems

Chairperson	Dr. Atef A. Elassal	U.S.A.
Co-chairperson	Dr. Mostafa Radwan	Netherlands
Secretary	Dr. Roop Malhotra	U.S.A.
Sessions:	G-2	

InterCommission WG II/III - Digital Photogrammetric Systems

Chairperson	Dr. Ian Dowman	United Kingdom
Co-chairperson	Dr. David McKeown	U.S.A.
Secretary	Zubbi Nwosu	United Kingdom
Sessions:	E-1, J-2	

Commission II Special Project - Upgrading Photogrammetric Instruments

Chairperson	Dr. Klaus Szangolies	Germany
Sessions:	D-1	

Important ISPRS Meetings

ISPRS Council Meeting	Sunday, June 5	9:00 - 6:00 PM
WG II/3 Business Meeting	Monday, June 6	1:30 - 3:00 PM
WG II/5 Business Meeting	Monday, June 6	3:30 - 5:00 PM
WG II/1 Business Meeting	Tuesday, June 7	8:30 - 10:00 AM
WG II/2 Business Meeting	Wednesday, June 8	8:30 - 10:00 AM
WG II/4 Business Meeting	Wednesday, June 8	8:30 - 10:00 AM
Business Meeting Commission II Executive Members	Wednesday, June 8	3:00 - 5:00 PM
InterCommission II/III Business Meeting	Thursday, June 9	8:30 - 10:00 AM
Commission II Special Project Business Meeting	Thursday, June 9	8:30 - 10:00 AM
Commission II Executive Meeting	Friday, June 10	2:00 - 4:00 PM

TABLE OF CONTENTS

Monday, June 6, 1994

TABLE DES MATIÈRES

Le lundi 6 juin 1994

Le vendredi 10 juin 1994

Séance L-1 - Groupe de travail II/1 - Technologies de cartographie en temps réel - Aspects algorithmiques

Plénière conjointe V - ISPRS/SIG 1994

APPLICATIONS OF GIS IN THE ENVIRONMENT

Merv Swan
Intera Technologies Ltd.

Abstract

Not available at time of printing

Applications environnementales des SIG

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif

SYNERGY OF PHOTOGRAMMETRY, REMOTE SENSING AND GIS – THE MOMS EXAMPLE

Dieter Fritsch

Institute of Photogrammetry, Stuttgart University, Keplerstrasse 11, D-70174 Stuttgart, Germany

CGIS/ISPRS Commission II Conference 1994, Plenary Paper

KEY WORDS: Modular Optical Multispectral/Stereo Scanner, digital terrain model generation, DTM integration in GIS, classification, Russian Space Station MIR

ABSTRACT

Digital spatial data acquisition and processing demand a high degree for automation and integration. Automation is necessary to process the huge amount of data captured by digital airborne and spaceborne sensors; integration allows further application of processed data in Geographic Information Systems (GIS).

Using the German development MOMS (Modular Optoelectronic Multispectral/Stereo Scanner) as an example – a digital spaceborne sensor for photogrammetric and remote sensing applications – the paper gives guidelines for further sensor developments and its data integration in GIS.

The MOMS sensor of its 2nd generation (MOMS02) is capable to capture three-fold stereo imagery within along-track movement, high resolution images up to 4.4m ground pixel size and multispectral images with 13.2m ground pixel size. It was flown onboard the Space Shuttle from April 26th till May 6th, 1993 during the 2nd German Spacelab Mission. In the meantime, first data are processed and the results give optimism for the synergy of photogrammetry and remote sensing.

The MOMS digital terrain models are integrated in GIS, furthermore thematic evaluations are also carried out and are linked to semantic data models. For this reason, MOMS data processing delivers for the first time a combination of terrain data and semantic data simultaneously to be stored in GIS. From May 1995 the MOMS sensor will be flown onboard the Russian Space Station MIR for a three years' term where besides stereo and multispectral data temporary aspects are of importance.

KURZFASSUNG

Die digitale Daterfassung und -verarbeitung erfordert eine hohen Automations- und Integrationsgrad. Automation ist die unabdingbare Voraussetzung zur Bearbeitung der großen Datenmengen, die durch flugzeug- oder satellitengetragene digitale Sensoren anfallen; die Integration erlaubt die Nutzung von prozessierten Daten in Geo-Informationssystemen (GIS).

Am Beispiel der deutschen Entwicklung MOMS – ein digitaler Erdbeobachtungssensor für photogrammetrische und thematische Anwendungen – gibt der vorliegende Beitrag Richtlinien für weitere Sensorentwicklungen und dessen Datenintegration in GIS.

MOMS02 ist in der Lage, Stereo-Bilddaten nach dem Pushbroom-Prinzip in drei Zeilen entlang der Sensorbewegung zu erzeugen, ebenso steht die Erfassung von hochauflösenden Bildern von bis zu 4.4m Bodenpixel im Vordergrund neben der gleichzeitigen Erzeugung von Mulitspektralbildern mit einer Auflösung von 13.2m. Dieser Sensor wurde während der 2. Deutschen Spacelab-Mission vom 26. April bis 5. Mai 1993 an Bord des Space Shuttle erfolgreich eingesetzt. Mittlerweile konnten erste Datensätze prozessiert werden; die Ergebnisse geben Optimismus für die Synergie von Photogrammetrie und Fernerkundung.

Die digitalen Geländemodelle (DGM) automatisch abgeleitet mit MOMS02-Daten werden in GIS integriert, des weiteren befinden sich thematische Auswertungen in der Bearbeitung, die mit thematischen GIS-Modellen zu verknüpfen sind. Aus diesem Grund liefert die Auswertung der MOMS02-Daten zum ersten Mal die Gelegenheit der simultanen Kombination von Geländegeometrie und -topographie abzuspeichern in GIS. Ab Mai 1995 wird der MOMS02-Sensor an Bord der russischen Raumstation MIR für etwa drei Jahre zur kontinuierlichen Datenerfassung eingesetzt, wobei neben der gleichzeitigen Erzeugung von Stereo- und Multispektralaufnahmen zeitliche Aspekte ebenso von Bedeutung sind.

La synergie photogrammétrie - télédétection - SIG

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif

1. INTRODUCTION

Integration of different disciplines has become a major task, what is already be acknowledged and under conceptual considerations and practical realizations. Photogrammetry with its main application field dealing with air surveys is on the one side a cost effective discipline, because a huge amount of object points can economically be delivered. Including digital image processing techniques the derivation of digital terrain models (DTM) is solved automatically (F. Ackermann/M. Hahn, 1991) without any operator control. On the other side restrictions of photogrammetry so far was analog image data acquisition. To overcome this ' defect ' aerial images are scanned and then ready for a digital workflow. Though digital sensors for aerial photogrammetry are under development (C. Thom/I. Jurvillier, 1993, O. Hofmann et al., 1993) it is not expected that such systems will operate in practice before the next millenium. Probably, the next 5-10 years aerial photogrammetry will operate hybrid, what means captured analog imagery is digitized in high-resolution scanners. Two products of digital photogrammetry are already generated in today's practice: DTM's and orthoimages (orthophotos).

Sensor development in remote sensing very soon concentrated on digital image acquisition. With the Landsat programme a series of digital spaceborne sensors was and is still in operation – further improvements aimed at a higher pixel resolution and stereo capability (for example Landsat, Spot).

In Germany since the midst of the seventies an Earth Oberservation Programme called MOMS is under development. It was put into operation onboard NASA's Space Transportation System (STS, Space Shuttle) three times. MOMS is an acronym for Modular Optical Multispectral/Stereo Scanner. What are the objectives behind this German Earth Observation Programme? The guidelines behind this programme have been worked out by the German Space Agency (DARA, Deutsche Agentur für Raumfahrtangelegenheiten) especially to

- establish and develop the technical and scientific basis for the utilization of space techniques
- acquire global, regional and local information on the earth surface as far as MOMS is concerned
- provide adequate technical systems and scientific validated methods for data processing, evaluation and interpretation
- demonstrate the applicability of earth observation from space
- support the transition of experimental systems into operational ones
- sponsor pilot programmes

As far as the MOMS programme is concerned the *milestones* of its developments are as follows. The programme started with first with investigations on the utilization of CCD line scanner imagery in 1977. At that time the Electro Optical Scanner (EOS) - today to be seen in The German Museum Munich - with one CCD line array was flown in an aircraft; further an engineering model of MOMS was built.

The first success of the MOMS sensor of its 1st generation – being indicated simply by MOMS01 – were the two flights with the NASA Space Shuttle about ten years ago. Besides the verification of the technology in space approximately 1000 scenes with an at that time unprecedented resolution of 20 m were collected. The data were evaluated in the frame of an international announcement of opportunity and a data exchange agreement with NASA.

Using the experience of MOMS01 camera design the second generation MOMS02 was developed. This sensor is quite different from MOMS01 because it fulfills the following requirements:

- three-fold stereo imagery
- along-track stereo imagery
- high resolution panchromatic imagery
- multispectral imagery
- combination of stereo and multispectral imagery

This sensor was flown onboard the second German Space Shuttle Mission (D2) from April 26 till May 6, 1993 with the objectives to generate automatically digital terrain models and its follow-up products with high accuracy, e.g. standard deviations in height less than 5 m, further to allow for data fusion of thematic mapping with digital terrain models and to integrate the results in Geographic Information Systems (GIS). At present the MOMS02 sensor is slightly modified to operate onboard the Russian Space Station MIR from 1995 to 1997 (onto segment PRIRODA). The objectives behind this German/Russian cooperation are as follows:

- first German/Russian cooperation in Earth Observation
- optimization of data evaluation algorithms for operational applications
- multisensor utilization
- development of a marketing strategy for earth observation data

But how will the resulting knowledge up to now and of MOMS02/PRIRODA be used for further developments? At present it is still an open question in Germany, although we had some discussions about a MOMS satellite to be operational towards the end of this millenium with full area coverage and a resolution on ground of 1 - 2 m.

2. MOMS02 TECHNICAL PARAMETERS

The main advantage of the MOMS02 sensor in comparison with existing earth oberservation satellites , for example LANDSAT TM and MSS as well as SPOT are as follows.

First MOMS02 has the highest resolution of a digital earth observing sensor we ever had. The high resolution panchromatic channel has about 4.4 m^2 ground pixel resolution considered from an altitude of 296 km. But besides that it offers for the first time along-track stereo capability which means the stereo-imagery geometry is overlapping during flight in one orbit contrary to the stereo capability of SPOT generated by data takes from different orbits. The MOMS02 stereo imagery concept can be seen in figure 1. A stereo image is generated by forward and backward looking lenses with a ground pixel resolution of about 13.2 m^2. These data are to be fusioned with the high resolution channel. This reason gives the optimism to generate digital terrain models of less than 5 m in height accuracy. The base line of a stereo scene is about 120 km with a swath width of 78 km, except the high resolution channel with a swath width of 37 km.

The MOMS02 camera was designed and built by DASA Munich (former MBB Munich) in close cooperation with DARA and the MOMS02/D2 science team with three principal investigators: Prof. F. Ackermann, Stuttgart University (photogrammetry), Prof. J. Bodechtel, University of Munich (thematics) and Prof. F. Lanzl, DLR Oberpfaffenhofen (combination of stereo and thematics). Both groups – stereo and thematics – are driven by co-investigators and a powerful team of 12 young scientists. The general tasks of the science team was besides the cooperation for camera design the software development to process the data taken by the D2 mission.

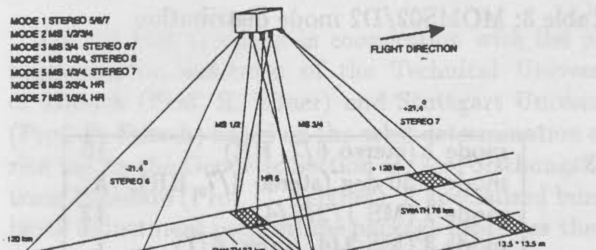


Fig.1: Along-track stereo capability

How does the technology of the MOMS02 camera look like? As indicated before the camera has altogether five lenses. The stereo module carries one vertically looking lens with focal distance of 660 *mm* for the high resolution channel and two tilted lenses with focal distance 237 *mm*. In figure 2 one can see that these three lenses are quite dominating. The other two lenses are used for multispectral imagery, each of it has a focal length of 220 *mm* and two CCD line arrays for different spectral recording characteristics located in the focal plane.

To summarize the MOMS technology the stere module consists of generally four Fairchild 191 CCD line arrays with each 6000 pixels/line whereby two line arrays have been merged each other for the high resolution panchromatic channel. The tilting of the backward and forward channel is about 21.43 degrees - all three line arrays are sensitive within the panchromatic spectral range of 520-760 nanometers and have a resolution in radiometry of 7 bit.

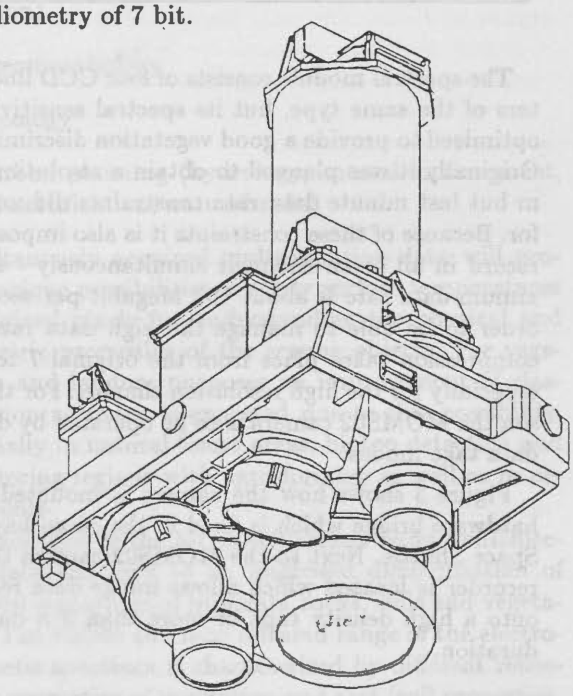


Fig. 2: MOMS02 optics module

Table 1: Stereo module parameters

four CCD line detectors pixel size 10x10μm	FAIRCHILD CCD 191
high resolution channel	12000 pixel/line (8500 are used)
backward channel	6000 pixel/line
foreward channel	6000 pixel/line
convergence angle	21.43 deg
spectral range	520 - 760 *nm*
focal length	HR: 660 *mm* BW/FW: 237 *mm*
radiometric resolution	7 bit
geometric resolution (296 *km* altitude)	HR: 4.4 *m* BW/FW: 13.2 *m*

Table 2: Multispectral module parameters

four CCD line detectors	FAIRCHILD CCD 191
pixel size 10x10 μm	
number of pixel	6000 pixel/line
spectral range	440 - 505 *nm*
	530 - 575 *nm*
	645 - 680 *nm*
	770 - 810 *nm*
focal length	220 *mm*
radiometric resolution	8 bit
geometric resolution	13.2 *m*
(296 *km* altitude)	

The spectral module consists of four CCD line detectors of the same type, but its spectral sensitivity was optimized to provide a good vegetation discrimination. Originally it was planned to obtain a resolution of 8.8 m but last minute data rate constraints did not allow for. Because of these constraints it is also impossible to record in all seven channels simultaneously - the maximum data rate is about 100 Megabit per second. In order to be able to manage the high data rates data compression takes place from the original 7 to 6 bits especially for the high resolution channel. For that reason the MOMS02 camera is to be operated by different data take modes.

Figure 3 shows how the camera is mounted onto a hardware bridge which is fixed in the cargo bay of the Space Shuttle. Next to the MOMS02 camera the data recorder is located which allows image data recording onto a high density tape of more than 5 *h* data take duration.

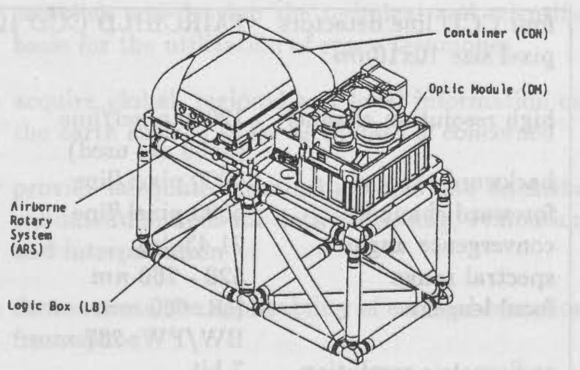


Fig. 3: MOMS02 camera and recorder mounted onto a hardware bridge

3. MOMS02/D2 MISSION

During the D2 mission last year image data could be recorded within a latitude of +-28.5 degrees. The weight of the Spacelab was about 6000 kilograms, therefore the Space Shuttle was not able to have a steeper inclination. The dotted lines in figure 4 indicate the ground tracks of the Space Shuttle, especially those where data takes could be taken. During the D2 mission the MOMS02 Science Team was in close contact with the shuttle to have control on the data takes dependent on the wheather conditions which were available in realtime in the German Space Operation Centre (GSOC) Oberpfaffenhofen (near Munich).

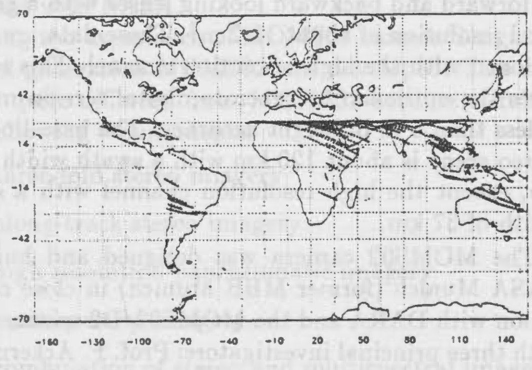


Fig 4: Ground tracks of MOMS02 data takes (from DLR 1993)

As indicated by table 3 altogether 48 data takes were made: 10 in normal stereo mode, 5 in an by 30 deg inclined stereo mode, 13 in combination of the high resolution panchromatic channel with the multispectral channels, and other combination as indicated. The total recording time exluding calibration was about 4 1/2 *h* with a total area coverage of approximately 7 Million km^2.

Table 3: MOMS02/D2 mode distribution

mode 1 (stereo 6/7, HR)	10
mode 1 30 deg (stereo 6/7, HR)	5
mode 2 (MS 1/2/3/4)	12
mode 3 (MS 3/4, stereo 6/7)	7
mode 4 (MS 1/3/4, stereo 6)	-
mode 5 (MS 1/3/4, stereo 7)	1
mode 6 (MS 2/3/4, HR)	13
mode 7 (MS 1/3/4, HR)	-

A quicklook of a scenery taken over the northern part of India is given by figure 5. In this image only every

5th pixel is used vor visualization.

Fig. 5: Quicklook of a MOMS02/D2 digital image

4. MOMS02 DATA PROCESSING

4.1 Photogrammetric objectives

The general scope of photogrammetry is the evaluation of the MOMS02 stereo-image data with regard to the task of topographic mapping, in particular to establish topographic databases on the basis of digital terrain models. Investigations are therefore made in a three-part programme (F. Ackermann, 1994):

- reconstruction of strip geometry and orientation
- automatic derivation of digital terrain models
- topographic mapping, setup of topographic databases

The 1st task is solved in cooperation with the photogrammetric institutes of the Technical University of Munich (Prof. H. Ebner) and Stuttgart University (Prof. D. Fritsch) based on the orbit determination carried out by the Geodetic Section of GeoForschungsZentrum Potsdam (Prof. C. Reigber). A specialized bundle block adjustment programme package processes the tie points found by image matching techniques.

The 2nd task is solved at the photogrammetric institute, Stuttgart University. Based on the experience to generate automatically digital terrain models in a hierarchical procedure, iconic and symbolic image pyramides are generated and processed to arrive at a regular grid DTM.

The 3rd task is solved at the photogrammetric institute, University of Hannover (Prof. G. Konecny). In this processing step, planimetric data are captured and stored in GIS. Also orthoimages are generated and compared with existing topographic databases. At this institute, also a special photogrammetric stereo workstation is under development being able to process and orientate three-line imagery.

4.2 Thematic objectives

The main investigations in thematic data processing will concentrate on the fields of (J. Bodechtel et al., 1994)

- land cover (vegetated areas: land mass, biomass estimation,agriculture, forestry, unvegetated areas: lithology, mineral exploration, tectonic investigations, pattern analysis and object recognition)
- geomorphology
- ecology
- urban planning (hydrology, coastal environment, desertification, natural risks)

Simultanously acquired multresolution data will provide unique possibilities to study surface temperatures and mixed pixels fur understanding the physical and geometric properties of the scanne objects. For vegetation and landuse purposes an improvement of classification accuracy is expected due to this possibility, especially in natural forest areas, biotop detection and monitoring regions with agro-forestry as well as in urban zones.

In addition to the MOMS02 narrow band multispectral data will lead to an improved differentiation of spectral signatures of minerals, rocks, soils and vegetation. The visible and near infrared range of the electromagnetic spectrum is characterised by different reflectance properties of vegetation and rock/soil properties. The spectral reflectance of minerals is maily influenced by the wings of charge transfer bands centred in the ultraviolet and by crystal field absorptions at longer wavelength, which are caused by transition elements. Furthermore, during proccessing of MOMS02/D2 data its suitability for atmospheric research (scattering, thickness of layers), hydrological and coastal studies, the detection and classification of instable slopes and others will be investigated.

5. MOMS02 ON PRIRODA

As already pointed out before the MOMS02 camera is right now under slight modification to be flown on the Russian Space Station MIR in May 1995. The orbital parameters of MIR are much better than the ones of the Space Shuttle flight last year. MIR has an attitude of about 350 - 400 *km* which allows data takes of about 60 of the earth. During MOMS02/P mission the number of operation modi is reduced to 3 – the stereo mode, the high resolution spectral mode and a combination of multispectral with one of the stereo channels. Because of the different figure in altitude the ground pixel resolution results in 5.2 - 6.1 *m* for the high resolution panchromatic channel and 15.9 - 18.2 *m* forthe lower resolution channels.

The camera with its data recorder will be mounted on top of the segment PRIRODA. One further component for this mission is a NAV-package allowing GPS positioning with an accuracy of about +- 5 *m* and attitude determination with an accuracy of +- 5 arcseconds. The data takes during that preparational mission should allow mapping and updating of GIS data bases. The operation scenario should be as follows: data takes will be made on demand by the users; the recorded data are transmitted to earth via parabolic antennae to Neustrelitz near Berlin and to Moscow. It is expected that this scenario can fulfill several requirements on earth observation data.

6. CONCLUSIONS

The MOMS programme of Germany is up to now very successful. About 30 Million DM have been spent for camera development and instrumentation – not too much in comparison with activities of the Landsat programme of The United States and the French Spot facilities. The camera itself has the highest ground pixel resolution of a digital space sensor ever flown; its along-track stereo capability allows for photogrammetric stereo restitution and automatic DTM generation. Furthermore, as shown by first data processing, the optimization of the width of multispectral bands leads to clear separation of vegetation, which could not be solved as efficient as desired in the recent past. The MOMS Science Team is very optimistic to obtain the accuracies predicted.

Integration of photogrammetric and remote sensing data in GIS is a necessity. Therefore, as shown by the MOMS example the cooperation of different experts in the disciplines of photogrammetry, remote sensing and geographic information systems has led to a development capable to manage geometric (stereo) and thematic (multispectral) data acquisition simultaneosly. The integration of the processed data in GIS can be done in different ways: a simple strategy is the layer oriented storage of different thematics – as far as DTM organization is concerned a more rigorous way is a total merge of planimetry with the DTM data. The MOMS programme is the optimum example of synergy; it is expected that during processing of the MOMS02/D2 data and MOMS02/P data further details for earth observation will be worked out which will have influence on further sensor design and permanent monitoring of the earth surface from space.

7. REFERENCES

Ackermann, F., Hahn, M., 1991: Image pyramids for digital photogrammetry. In: Digital Photogrammetric Systems, Eds. H. Ebner, D. Fritsch, C. Heipke, Wichmann, Karlsruhe, pp. 43-58.

Ackermann, F. 1994. The photogrammetric research program with MOMS-02/D2 stereo-image data from space. In: Proceedings Commission 3, XX. FIG Congress, Melbourne, Australia, pp. 304.1/1-10.

Bodechtel, J., Lörcher, G., Sommer, S., (1994). The MOMS-02/D2 Mission – Overview on the thematic evaluation. In: Proceedings Commission 3, XX. FIG Congress, Melbourne, Australia, pp. 304.3/1-9.

Edwards, G., 1991. Remote sensing image analysis and geographic information systems: laying the groundwork for total integration. In: The Integration of Remote Sensing and Geographic Information Systems, Ed. J.L. Star, American Soc. for Photogrammetry and Remote Sensing, pp. 21-31.

Ehlers, M., Greenlee, D.D., Star, J.L., Smith, T.R., 1991. Integration of remote sensing and GIS. Photogr. Engin. Rem. Sens. (PERS), 57, pp. 669-675.

Ehlers, M., 1993. Developments in remote sensing: techniques and applications. In: Multisource Data Integration in Remote Sensing for Land Inventory Applications, Eds. M. Molenaar, L. Janssen, H. van Leeuwen, Dept. Land Surveying and Remote Sensing, Wageningen Agricultural University, pp. 55-64.

Fritsch, D., 1991. Raumbezogene Informationssysteme und digitale Geländemodelle. Deutsche Geodätische Kommission, Reihe C, Nr. 369, München.

Fritsch, D., 1993. Photogrammetry and geographic information systems – evolution instead of revolution. In: Photogrammetric Week'93, Eds. D. Fritsch, D. Hobbie, Wichmann, Karlsruhe, pp. 3-10.

Hofmann, O., Kaltenecker, A., Müller, F. (1993): Das flugzeuggestützte, digitale Dreizeilenaufnahme- und Auswertesystem DPA - erste Erprobungsergebnisse. In: Photogrammetric Week'93, Eds. D. Fritsch, D. Hobbie, Wichmann, Karlsruhe, pp. 97-107.

Sahm, P.R., Keller, M.H., Schiewe, B., 1993. Research program of the German Spacelab Mission D-2. DLR, Cologne, 558 p.

Seige, P. (1993). MOMS – A contribution to high resolution multispectral and stereoscopic earth observation from space. In: Photogrammetric Week'93, Eds. D. Fritsch, D. Hobbie, Wichmann, Karlsruhe, pp. 109-120.

Thom, C., Jurvillier, I. (1993): Experiences with a digital aerial camera at Institut Geographique National (France). In: Photogrammetric Week'93, Eds. D. Fritsch, D. Hobbie, Wichmann, Karlsruhe, pp. 73-79.

Zilger, J., Seige, P. (1994). The Modular Optoelectronic Multspectral Stereo Scanner MOMS – A German Program for Earth Observation. In: Proceedings Commission 3, XX. FIG Congress, Melbourne, Australia, pp. 304.2/1-6.

SAR IMAGE LOOK DIRECTION BIAS CORRECTION USING WAVELET TRANSFORM

Wooil M. Moon[1], J.S. Won[2], and X.G. Miao[1]

1. Geophysics, The University of Manitoba, Winnipeg, Canada R3T 2N2
(wmoon@cc.umanitoba.ca)

2. Marine Geophysics, Korea Ocean Research and Development Institute
P.O. Box 29, Ansan City, Ansan, Kyunggi-do, Korea
(jswon@sari.kordi.re.kr)

ABSTRACT

Look direction bias in a single look SAR image has been a problem in the geological application of SAR data, particularly in the relatively flat Precambrian shield terrain. This paper investigates digital processing techniques for airborne and space-borne SAR image data integration and for correction of look direction bias problem. The two important approaches for reducing the look direction bias problem are principal component analysis (PCA) and Wavelet transform (WT) techniques.

In this research we investigated and tested the techniques with ERS-1, JERS-1 and CCRS's airborne SAR data sets. The PCA technique has been proved to be very effective in various remote sensing applications. The PCA technique, however, requires three (or at least two and one auxiliary) data sets for rendition of the properly corrected final image. The Wavelet transform approach utilizes the property which decomposes image data into approximated image (low frequency) and detailed image (high frequency). The Wavelet transform approach is more efficient and robust in enhancing the fine details of the multiple SAR image data without sacrificing the original image resolution.

KEY WORDS : SAR, Wavelet transform, Geological remote sensing, Look direction bias

IMAGERIE INTÉGRÉE DES DONNÉES DES SATELLITES RADAR À SYNTHÈSE D'OUVERTURE JERS-1 ET ERS-1 DE RECHERCHE SUR L'ACTIVITÉ SISMIQUE DANS LA RÉGION DE NAHANNI, TERRITOIRES DU NORD-OUEST, CANADA

Résumé

La région de Nahanni, dans les Territoires du Nord-Ouest, est une région reculée à forte activité sismique. Dans le cadre d'un projet coopératif de recherche, notre équipe a étudié les données des satellites de télédétection à radar à synthèse d'ouverture JERS-1 et ERS-1, et des satellites SPOT et Landsat, en plus d'autres données géophysiques portant notamment sur l'épicentre des chocs principaux et des répliques des secousses sismiques. Il y a eu dernièrement plusieurs importants tremblements de terre près de la rivière Nahanni nord, y compris ceux du 5 octobre 1985 (magnitude de 6,6), du 23 décembre 1985 (magnitude de 6,9) et du 25 mars 1988 (magnitude de 6,2) Ces tremblements de terre se sont tous produits dans un plateau relativement peu déformé et de faible superficie, soit la plaine du Mackenzie, dans la ceinture orogénique de l'avant-pays, formée le long de la cordillère nord-est pendant la phase colombienne ou laramienne. L'une des principales difficultés que présente l'étude du cadre géologique et tectonique de la région de l'épicentre de Nahanni consiste en l'absence de données géologiques et géophysiques complètes, sans doute imputable à l'éloignement et aux difficultés logistiques et d'accessibilité qui en découlent. Dans la présente étude, nous avons tenté d'utiliser les données numériques et optiques des images obtenues par radars à synthèse d'ouverture fonctionnant en bande L et en bande C (Landsat et SPOT) pour compléter d'autres données existantes. L'étude détaillée des caractéristiques structurales de la configuration de surface, ainsi que des images distinctes et des images composites finales intégrées de la zone de l'épicentre, a révélé la présence de nouvelles caractéristiques structurales qui coupent la principale zone d'orientation préférentielle nord-ouest de la faille de compression de la chaîne Iverson, ainsi que des structures connexes dans la zone de l'épicentre du tremblement de terre.

INTEGRATED IMAGING OF JERS-1 AND ERS-1 SAR DATA FOR EARTHQUAKE TECTONIC INVESTIGATION OF THE NAHANNI EARTHQUAKE AREA IN NORTHWEST TERRITORIES, CANADA

W.M. Moon[1], V. Singhroy[2], Roy Kuoda[3], Maurice Lamontagn[4] and Y. Yamaguchi[3]

1. Geophysics, The University of Manitoba, Winnipeg, Canada R3T 2N2
2. Applications division, Canada Center for Remote Sensing, Ottawa, Canada K1A 0Y7
3. Geological Survey of Japan, Tsukuba, Ibaraki, 305 Japan
4. Seismology, Geological Survey of Canada, Ottawa, Canada K1A 0Y3

ABSTRACT

The Nahanni region, Northwest Territories, Canada is a remote area and has been seismically very active. As a collaborative research project, the investigation team acquired JERS-1 SAR, ERS-1 SAR, SPOT and Landsat data in addition other geophysical data such as earthquake epicenter data for main shocks and after-shocks. Recently there have been several large earthquakes near the North Nahanni River including the ones on 5 October 1985 (Ms 6.6), 23 December 1985 (Ms 6.9) and 25 March 1988 (Ms 6.2). The earthquakes occurred within a small relatively undeformed plateau, the Mackenzie Plain, in the Foreland Fold Belt created along the northeastern Cordillera during the Columbia or Laramide Orogeny. One of the problems in studying the geological and tectonic setting of the Nahanni epicentral region has been lack of comprehensive geological and geophysical data, perhaps due to remote isolation and logistic difficulties in accessing the study area. In this study an attempt was made to use space-borne L- and C-band SAR digital image and optical (Landsat and SPOT) image data to compliment other available information. Detailed investigation of the surface structural features of individual as well as the final integrated composite images over the epicentral area indicates that there are several new structural features intersecting the major northwest trending Iverson thrust fault and accompanying structures at the earthquake epicentral region.

KEY WORDS : SAR, Geological Remote Sensing, Earthquake, Integration

CORRECTION DE JUSTESSE DE LA DIRECTION DE VISÉE EN IMAGERIE DE RADAR À SYNTHÈSE D'OUVERTURE, AU MOYEN DE LA TRANSFORMÉE D'ONDELETTES

Résumé

La justesse de la direction de visée, en imagerie par radar à synthèse d'ouverture à visée unique, laisse beaucoup à désirer en ce qui a trait aux applications géologiques de ce mode de télédétection, plus particulièrement lorsqu'il s'agit de cartographier un sol de type précambrien relativement plat. Le présent document examine les possibilités des techniques numériques de traitement en matière d'intégration des données d'images obtenues par satellite de télédétection à radar à synthèse d'ouverture et en ce qui a trait à la correction de ce défaut de justesse de la direction de visée. L'analyse en composantes principales et la transformée d'ondelettes sont les deux techniques essentielles employées pour tenter d'atténuer ce problème.

Ces techniques ont été mises à l'essai sur les ensembles de données de télédétection obtenues par satellites à radar à synthèse d'ouverture ERS-1 et JERS-1 et à radar à synthèse d'ouverture aéroporté du Centre canadien de télédétection. La technique d'analyse en composantes principales s'est avérée particulièrement efficace dans diverses applications de télédétection. Cette technique a cependant nécessité l'emploi de trois ensembles de données (ou d'au moins deux ensembles de données et d'un ensemble auxiliaire) pour obtenir une image finale adéquatement corrigée. Quant à la technique de transformée d'ondelettes, qui décompose une image donnée en une image approchée (basse fréquence) et en une autre image détaillée (haute fréquence), elle s'est révélée particulièrement efficace et fiable pour rendre la finesse des détails d'images multiples de radar à synthèse d'ouverture.

Dunes: A Synthetic Aperture Radar (SAR) View

Dan G. Blumberg and Ronald Greeley
Department of Geology
Arizona State University
Tempe, Arizona 85287-1404
e-mail address: Blumberg@ASUIP2.DNET.NASA.GOV

Synthetic Aperture Radar (SAR) provides a unique perspective of Earth's surface and processes involved in its formation. Features formed by windblown sand and dust are abundant in deserts and are sensitive to climate and climate change. The JPL SAR (AIRSAR) was used to study surfaces in the Mojave. Results show that SAR provide information both on form and process as they pertain to windblown sand. High resolution SAR images can decipher dune types (i.e., star, linear, transverse dunes, etc.), and can provide an indication of surface roughness -- a key parameter in modeling aeolian activity.

We show SAR signatures of dune types including: linear, transverse, and star dunes. The star dunes are easiest to identify in radar images due to a distinct blooming signature. Linear and transverse dunes can be identified due to their speckle return and dark interdunal surface. Interestingly, we noticed in radar images of the Stovepipe Wells dune field that even when the illumination is parallel to the crest of the dunes they are visible.

Generally, calibrated SAR data can provide a useful tool in studying dryland environments. The anticipated launches of the Shuttle Radar Laboratory (SRL) will provide many scenes of wind swept regions on Earth. We anticipate that we will show some results from the SRL mission.

Dunes : visualisation par radar à synthèse d'ouverture (RSO)

Résumé

Le radar à synthèse d'ouverture (RSO) fournit une perspective unique de la surface de la terre et des processus qui la façonnent. Dans les déserts, le sable soufflé par le vent sculpte de très nombreuses formations qui sont influencées par le climat et ses changements. Le RSO du JPL (AIRSAR) a servi à étudier la surface du désert de Mojave. Les résultats ont montré que le RSO fournit de l'information touchant les formes et les processus liés au sable soufflé par le vent. Les images à haute résolution du RSO permettent de distinguer les types de dunes (par exemple, à arête radiale, linéaire, transversale, etc.) et donnent une indication de la rugosité des surfaces, paramètre clé de la modélisation de l'activité éolienne.

Nous présentons des signatures RSO de types de dunes (linéaires, transversales, à arête radiale, etc.). Les dunes à arêtes radiales sont les plus faciles à identifier dans les images radar grâce à leurs flous caractéristiques. Les dunes transversales et linéaires peuvent être repérées en raison de leurs tachetures et de leurs surfaces interdunaires plus sombres. Il est intéressant de remarquer que, dans les images radar de la zone où se trouvent les dunes de Stovepipe Wells, celles-ci sont visibles, même lorsque leur arête est parallèle au sens de l'illumination.

Généralement, les données étalonnées d'un RSO constituent un outil précieux pour l'étude des régions arides. Les vols prévus du Shuttle Radar Laboratory (SRL) permettront d'obtenir de nombreuses images des régions de la planète qui sont balayées par le vent. Nous prévoyons présenter certains résultats de ces missions.

Digital Topographic Data from Topsar Radar Interferometry: Fernandina Volcano Galapagos Islands

Peter J. Mouginis-Mark,

University of Hawaii, U.S.A.

ABSTRACT

In May 1993, the TOPSAR airborne Interferometric radar collected a digital topographic data set for Volcan Fernandina, Galapagos Islands. These data have a spatial resolution of 10 meters, and a vertical accuracy of about 2 meters. In addition to the topographic measurements simultaneous quad-pol L-band (24 cm) and P-band (70-cm) SAR data were collected. Almost the entire island (about 550 square kilometers) was mapped during a series of over flights by the radar, which is flown on the NASA/JPL DC-8. To help interpret these TOPSAR data, we have assembled a data base that includes Landsat Thematic Mapper, panchromatic SPOT, and JERS-1 OPS and SAR images. Furthermore, the island represents one of our Space Shuttle Radar (SIR-C/X-SAR) "Super-Sites", which will be imaged during flights planned for April and August 1994.

The resultant topographic map is a unique set for Volcan Fernandina, which has previously never been completely mapped topographically due to cloud cover. The TOPSAR data serve as both a test case for the validation of the radar interferometric measurements, and for the analysis of this infrequently visited basaltic shield volcano. For example, the production of shaded relief and slope maps from TOPSAR data, and their comparison with the SPOT and TM images, has proven to be the most useful analysis method to date. While lava flows can be mapped from the optical data, shaded relief maps that simulated diverse lighting geometries aid the identification of cinder cones and show that there are no prominent rift zones on the volcano that indicate structural control on the distribution of intrusions. Maximum slopes exceed 30-degrees on the upper flanks of the valcano, and pose major problems for current models for the internal structure of the volcano's summit area.

This paper will review the utility of TOPSAR data for volcano logical studies, and will include an analysis of the "error maps" which describe the quality of the phase correlation generated by the radar on a pixel-by-pixel basis. TOPSAR data were also collected for the adjacent Isabela Island (about 4,500 square kilometers), and an up-date for the on-going analysis of these data will also be presented.

Données topographiques numériques obtenues au moyen du radar interférométrique Topsar : le volcan Fernandina des îles Galapagos

RÉSUMÉ

En mai 1993, le radar interférométrique aéroporté TOPSAR a recueilli un ensemble de données topographiques numériques sur le volcan Fernandina, aux îles Galapagos. Ces données sont d'une résolution spatiale de 10 m et d'une précision verticale d'environ 2 m. En plus des mesures topographiques, on a recueilli simultanément des données au moyen d'un RSO quadripolaire à bande L (24 cm) et à bande P (70 cm). Une série de passages de l'avion porteur du radar, qui était monté à bord d'un DC-8 du JPL de la NASA, a permis de cartographier presque toute la surface de l'île (environ 550 km^2). Pour aider à interpréter cette information, nous avons monté une base de données qui comprend des images panchromatiques SPOT et d'autres images obtenues au moyen d'un appareil de cartographie thématique du Landsat, d'un RSO et du satellite d'observation JERS-1. De plus, cette île constitue l'un des «super-sites» que doit explorer notre radar (SIR-C/X-SAR), qui sera utilisé au cours des prochaines missions de la navette spatiale; des images seront prises pendant les vols prévus pour avril et août 1994.

La carte topographique qui doit en résulter constitue un ensemble de données unique sur le volcan Fernandina, dont la couverture nuageuse a jusqu'à présent empêché la production d'une carte topographique complète. Les données du TOPSAR ont servi de jeu d'essai pour la validation des mesures prises par radar interférométrique, ainsi que pour l'analyse de ce volcan en bouclier basaltique mal connu. Par exemple, la méthode d'analyse suivante s'est révélée être la meilleure à ce jour : la production de cartes topographiques au relief par ombres portées et de cartes de pentes, et la comparaison de ces documents avec des images SPOT et TM. Bien qu'il soit possible de cartographier des coulées de lave, les cartes au relief par ombres portées qui simulaient les diverses géométries d'éclairement aident à identifier les cônes de scories et à déterminer l'absence de zones de fracture saillantes sur le volcan, ce qui révèle le contrôle structural sur la distribution des intrusions. L'inclinaison maximale dépasse 30° sur les flancs supérieurs du volcan; cela pose des problèmes épineux en ce qui a trait aux modèles actuels de représentation de la structure interne de la zone du sommet.

L'article étudie l'utilité des données du TOPSAR pour les études formelles sur les volcans. Il comprend également une analyse des «cartes d'erreurs» qui décrivent, pixel par pixel, la qualité de la corrélation de phases des données obtenues par radar. Cet article contient enfin une mise à jour de l'analyse courante des données du TOPSAR qui ont aussi été recueillies sur l'île Isabella (environ 4 500 km^2 de superficie), qui est située à proximité.

Earth Observation Data Management at the Canada Centre for Remote Sensing

R. Boudreau[1], A. Buffam[2] and J. Guenette[2]

[1]*Canada Centre for Remote Sensing, 588 Booth Street, Ottawa, Ontario, Canada, K1A 0Y7*
[2]*Prologic Services Limited, 75 Albert Street, Suite 206, Ottawa, Ontario, K1P 5E7*

ABSTRACT

The Canada Centre for Remote Sensing (CCRS) has operated an information system, called GCNet, since 1990. This system is being used regularly by the Canadian and international resource managers, global change researchers and scientists in order to locate earth observation data sets in Canada or abroad. This paper will present an overview of the GCNet system capabilities, including access to the International Directory Network, the QUERY Image Inventory and the PlaNet bulletin board. Statistics on the utilization of GCNet will also be presented. Future developments in the areas of data management at CCRS will also be examined: on-line access to digital browse; international catalogue interoperability and links with other projects such as the Land Information Network for Canada (LINC) developed by the Surveys, Mapping and Remote Sensing Sector (SMRSS) of Natural Resources Canada, and the Canadian Earth Observation Network (CEONet).

RÉSUMÉ

Le Centre canadien de télédétection (CCT) exploite le système d'information GCNet depuis 1990. Ce système est utilisé régulièrement par les gestionnaires des ressources naturelles, les chercheurs intéressés par les transformations à l'échelle du globe et autres scientifiques canadiens et à l'étranger afin de localiser des jeux de donnés au Canada ou à l'étranger. Cet article présente un survol des capacités de GCNet: Réseau de répertoire international, l'inventaire d'images QUERY du CCT et le babillard électronique PlaNet ainsi que les statistiques d'utilisation du système depuis son installation. Enfin, on se penchera sur quelques développements à venir au CCT dans le domaine de la gestion des données d'observation terrestre: accès en ligne aux images minutes, interoperabilité entre les catalogues internationaux et les liens avec d'autres projets important tels le système Canada sous réseau de télécommunication (CARTE) développé par le Secteur des levés, de la cartographie et de la télédétection (SLCT) de Ressources Naturelles Canada et le Canadian Earth Observation Network (CEONet).

KEYWORDS: Global Change, Information System, Catalogue Service, Directory Service

1. INTRODUCTION

Since the launch, in 1972, of the first land remote sensing satellite, Landsat 1, large quantities of data have been acquired around the world. Although archiving has not always been uniform, many images have nevertheless become available to users from various groups including resource managers, scientists, students and the public.

When the number of images was relatively small, computer technology was not widespread and image browsing could be carried out at few locations using microfiche, the cataloguing and access systems supporting satellite data archives were quite primitive. They consisted essentially of a computer printout of the existing archived scenes and accompanying microfiche. From the viewpoint of the user, both were somewhat out of date because of the delays in updating, printing and processing of new acquisitions. More importantly, since few remote terminals existed and the database access software was not available, the access to information about archive data was rather limited.

As the volume of data increased and with growing interest in change detection and assessment, the need for more sophisticated tools became more apparent. GCNet is a system developed at CCRS in 1990 [1,2] in response to this requirement to provide users better access to remote sensing and other data required by Global Change scientists.

This article provides an overview of the services offered by GCNet in support of earth observation data management at CCRS; presents some statistics about GCNet utilization and looks at future developments and services to be offered by GCNet.

2. OVERVIEW OF GCNET

The GCNet system is developed around a Digital micro-VAX 3100 running under the VAX/VMS operating system. It is accessible 24 hours a day via Direct Dial, DATAPAC, Internet and NSI-DECnet.

The services provided by GCNet are the following: Directory, Inventory, Bulletin Board and a Remote-Sensing Bibliographic service (RESORS). There are no usage fees, except for RESORS, but the user may have to pay long distances charges where it applies. No special Username/Password is required with the exception of RESORS again.

• Directory Service

The GCNet Directory provides a single point of contact for Canadian Global Change scientists and other users who want to obtain information about CCRS and other Canadian data holdings. It also provides direct network links to international directories and inventories

The GCNet directory function is provided through the International Directory Network (IDN) Master Directory (MD) developed by NASA [3,4]. The MD data base contains descriptive information for over 2000 datasets in various countries. These datasets are organized by Disciplines and Sub-disciplines. Figures 1 and 2 give a breakdown of the datasets among them. The data base can be searched by numerous criteria such as parameter keywords, geographic location, and time. The MD data base implementation follows an internationally accepted standard [5]. The fields stored in the MD data base for each dataset are:

- Directory entry identifier
- Directory entry title
- Start and Stop dates
- Sensor name, Source name
- Investigator, Technical contact
- Data centre (name, contact person)
- Dataset identifier
- Originating centre, Campaign or Project name
- Storage medium
- Parameter measured
- Discipline keywords, Location keywords, General keywords
- Dataset Coverage
- Revision date, Science review date, Future review date
- Reference
- Summary

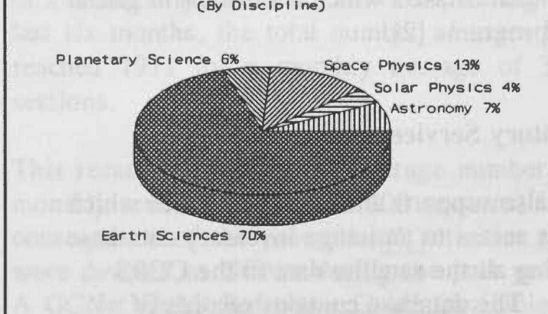


Figure 1 Distribution of the datasets in the IDN Master Directory among the five supported disciplines.

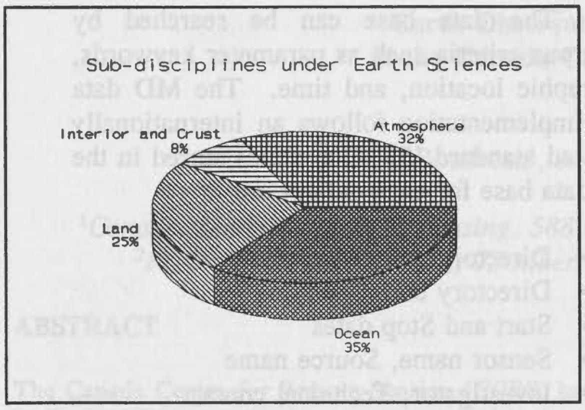


Figure 2 Distribution of the Earth Sciences datasets among the sub-disciplines Land, Atmosphere, Ocean and Interior & Crust.

When a dataset of interest has been identified the user can automatically connect to data systems and catalogues in Canada and around the world.

The development of the *Directory* service is based on *catalog interoperability* concepts which have been developed by the Committee on Earth Observation Satellites (CEOS) and widely endorsed by the international science community. This compatibility ensures that GCNet will provide Canadian users access to international datasets which are critical to global change programs [2].

• Inventory Service

GCNet also supports an *Inventory* service which provides access to an image inventory data base describing all the satellite data in the CCRS archive. The database contains records of approximately 1 million scenes which have been archived at CCRS since 1972. All future satellite data archived by CCRS, such as RADARSAT data processed on CCRS FASTSCAN [6] system, will be added to this data base.

The GCNet Inventory Service is provided by the CCRS QUERY program. QUERY provides the capability to search the Image Inventory data base which describes the CCRS data archive holdings. This capability is essential in view of the large volume of data presently available (Table 1).

Table 1 CCRS satellite data archive content for the period between July 1972 and February 1994.

MISSION	Scenes
Landsat MSS	387,000
Landsat TM	106,000
SPOT (Multispectral)	135,000
SPOT (Panchromatic)	316,000
MOS MESSR	31,000
SEASAT SAR	462
ERS-1 SAR	107,000
TOTAL	976,174

QUERY provides the following image inventory search capabilities:

i) Selection of geographical area either by: Lat/Long rectangular box or Track/Frame rectangular box.

ii) Selection of scenes based on any logical combination of the following fields:
- satellite identifier
- sensor identifier
- sensor mode (if applicable)
- view angle (if applicable)
- acquisition date
- cloud cover

iii) Multi-inventory searches: i.e. the capability to do searches in many satellite/sensor databases from a single user session.

For ease of consultation, the QUERY search results, in a tabular report format, can be e-mailed to the user workstation via Internet.

• **Bulletin Board**

The Bulletin Board service is provided through the CCRS PlaNet bulletin board system. This service is subdivided into three main sections: SMRSS Products & Services, CCRS Bulletin Board and Special Interest Groups.

The *SMRSS Products & Services* section allows users to view a directory of product descriptions, single and group prices and the various contacts in Canada for ordering analog and digital products.

The *CCRS Bulletin Board* provides a noticeboard, message area and news items. Information concerning upcoming conferences in remote sensing, CCRS projects and publications can be found under that topic. In addition, information about the CCRS Technology Enhancement Program, training and development can be found there.

The *Special Interest Groups* section contains information about some remote sensing activities at the provincial and territorial level, remote sensing company profiles and a list of educational institutes in remote sensing applications and information on the Canadian Remote Sensing Society.

A file transfer are allows users equipped with a PC and modem to download sample image files and a variety of CCRS publications onto their system.

• **RESORS**

RESORS is an on-line bibliographic database that provides rapid and precise access to information on the technologies and applications of remote sensing world-wide. It is internationally recognized as the leading bibliographic information source in remote sensing. It is owned and operated by the Canada Centre for Remote Sensing.

The RESORS database contains over 89,000 references that have been indexed by experienced subject specialists over the last two decades.

RESORS is extensively used by organizations and individuals throughout the world. It serves remote sensing scientists, engineers, researchers, educators and students. Governments, corporations and academic institutions all make substantial use of the system and people from virtually every country in the world have accessed RESORS.

There is an annual subscription cost for this retrieval service. Other products such as training manuals, slide sets, acronym lists and glossaries are also available.

3. GCNet USAGE STATISTICS

GCNet has been made available on November 1990. Since October 1992, we have monitored usage statistics. Table 2 shows the number of monthly sessions from October 1992 until March 1994 inclusively. For the first twelve months (Oct 1992 - Sep 1993), 2614 sessions were recorded on GCNet. This represents an average of 217 monthly sessions. However, during the last six months, the total number of users has reached 1971 for a monthly average of 328 sessions.

This recent increase in the average number of monthly sessions (more that 100 sessions) is a consequence of the promotional activities that were devoted to GCNet during the past years. A GCNet Quick Reference Guide (available in both English and French languages) has been printed and distributed to a wide audience via a few national journals/newsletters devoted to remote sensing. Many presentations have been made at national and even international conferences. Some advertisement was also made in remote sensing journals and contacts were established with several Canadian institutions involved in Global Change/Remote Sensing activities. Finally, if users have specific questions on how to use the services on GCNet they can get in touch with User Assistance. Comments and suggestions from users are also welcomed. In fact, many recent upgrades to GCNet were made in response from such comments and/or suggestions.

Table 2 Number of GCNet sessions since September 1992.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
1992										232	275	189	**696**
1993	155	177	230	232	194	224	231	224	251	230	368	294	**2810**
1994	411	303	365										**1079**

An interesting statistic concerns the way users connect to GCNet. Currently, there are five possible ways to access it: Internet, Datapac, NSI-DECnet, Dial-Up and Direct line (for CCRS users only). Table 3 looks at the distribution of GCNet sessions among the five connection paths. The figures are shows for the months of January 1993 and January 1994.

Table 3 This table illustrates the change during the last year in the way the users access GCNet. Internet is clearly emerging as the preferred access route.

Access via	Jan 93	Jan 94
Internet	17%	68%
Datapac	29%	10%
Dial Up	17%	6%
NSI-DECnet	8%	6%
Direct line	29%	10%

The number of Internet sessions has increased by a factor of four while the usage of the other connection paths has decreased by a factor of 3 (with the exception of NSI-DECnet). In fact, Figure 3 shows that this increase in the number of Internet sessions is not a random phenomena but is rather part of a trend related to the widespread and ever-increasing availability of Internet.

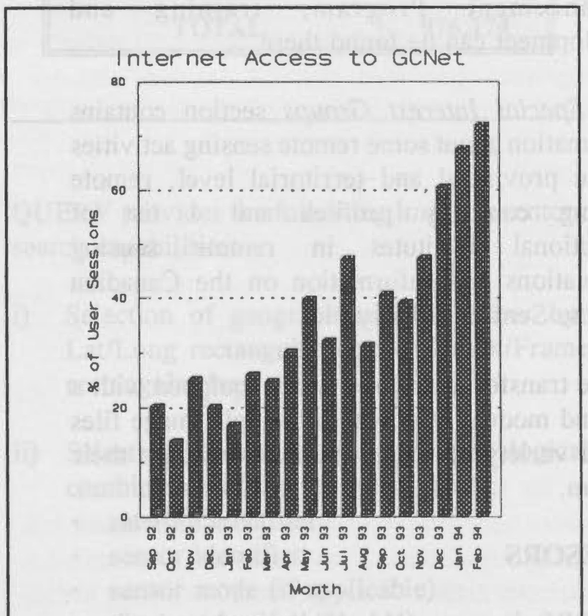


Figure 3 User access to GCNet via Internet. During the last year, the percentage of users accessing GCNet via Internet has quadrupled.

In order to take advantage of this, there is currently work going on to provide our users access to GCNet via Internet tools such as Gopher and World Wide Web. A Gopher server should be installed within the next year.

4. FUTURE DEVELOPMENTS

Global Change programs are *global* in scope, hence they often require data held in widely dispersed national inventories or in inventories of other countries. Much effort is now being directed toward the development of standards in the areas of directories, inventories and communication between remote directories and inventories. Groups actively pursuing these standards include: the NASA Catalog Interoperability Working Group (which has representatives from several United States agencies and other organizations such as the European Space Agency and CCRS), the Committee on Earth Observation Satellites (CEOS) Catalog Subgroup (which has representatives from many countries with active space programs) and the Earth Observation System (EOS).

An interesting initiative called Catalogue Interoperability Experiment (CINTEX), sponsored by CEOS, involves participants from United States and Europe, and has been under way for about two years now. The CINTEX participants are smoothing away the technical difficulties arising when trying to operate heterogeneous and distributed systems in a client-server environment.

A demonstration on catalogue interoperability was done at the last CEOS Catalogue Subgroup meeting in Ottawa in February 1994. An EOS catalogue search client demonstrated the capability to retrieve catalogue information from servers located in the United States and Europe. The communication process between the various systems involved is handled by the Internet network. Further work is going on to provide users with an interface as seamless and robust as possible.

CCRS feels that its users would benefit from such an information infrastructure and participation in CINTEX is planned. Initial plans are to have the Landsat and SPOT catalogues made accessible to the CINTEX community.

GCNet will also be tied in to other important projects such as Land Information Network for Canada (LINC) and Canadian Earth Observation Network (CEONet). These will allow access to on-line databases of products available from the SMRSS and from other Canadian earth observation agencies.

4. CONCLUSION

So far, the user response to GCNet has been very favourable. The usage of the various services offered keeps growing every year and user's comments and suggestions have been very valuable in upgrading the system such that it better fulfils their needs. Further developments will be done in order to make better use of the services available on Internet (Gopher, WWW).

CCRS is also actively participating in some international activities in the fields of catalogue information exchange, such as the International Directory Network, to ensure that Canadian Global Change users can access the data they require in a timely and coherent fashion. It is also important that the standards adopted by these groups meet our needs and that GCNet functions according to those standards.

5. REFERENCES

[1] Fisher, T.A. and R. Boudreau, 1991. "Data Management for Global Change". Proceedings of the 1991 Canadian Conference on GIS, March 18-22, Ottawa, Ontario, pp. 138-144.

[2] Fisher, T.A., J. Cihlar and R. Boudreau, 1991. "Global Change Network (GCNet". Proceedings of the 14th Canadian Symposium on Remote Sensing, Calgary, Alberta, May 6-10, pp. 153-157.

[3] National Space Science Data Centre, 1989. "The NASA Master Directory, Quick

Reference Guide", Goddard Space Flight Centre, Publication 89-19.

[4] Thieman, J., 1989. "NASA Master Directory, First Step to Data", Information Systems Newsletter, Goddard Space Flight Centre, p. 12-14.

[5] National Space Science Data Centre, 1989. "Directory Interchange Format Manual", Goddard Space Flight Centre, Publication 89-24.

[6] D'Iorio, Marc A. and Trevor Butlin, 1993. "The Canadian Radarsat FASTSCAN System", IGARSS'93 Symposium, Volume II, Tokyo, Japan, August 18-21, pp.568-570.

Gestion de données d'observation de la Terre au Centre canadien de télédétection

Résumé

Le Centre canadien de télédétection (CCT) exploite depuis 1990 un système d'information nommé GCNet. Ce système est utilisé régulièrement au Canada et à l'échelle mondiale par des gestionnaires de ressources, des chercheurs dans le domaine de la transformation du globe et des scientifiques en vue de repérer des ensembles de données d'observation de la Terre regroupées au Canada ou à l'étranger.

La présente communication donne une vue d'ensemble des possibilités du système GCNet, notamment l'accès à l'International Directory Network, au QUERY Image Inventory et au bulletin électronique PlaNet. Elle fournit également des statistiques sur l'utilisation du GCNet.

On y examine aussi l'expansion des services au CCT dans les domaines de la gestion des données : accès direct à la lecture d'information numérique; interopérabilité du catalogue international et liaisons avec d'autres projets tels que le CARTE (Canada sous réseau de télécommunication électronique) du Secteur des levés, de la cartographie et de la télédétection et le Réseau canadien d'observation de la Terre (CEONet).

Canadian Earth Observation Network (CEONet)

T. Fisher, N. Denyer
Canada Centre for Remote Sensing, Natural Resources Canada
P. Price
Canadian Space Agency, Industry Canada
H. Edel
Department of Fisheries and Oceans, Canada
L. Stirling
Atmospheric Environment Service, Environment Canada

Abstract

Earth observation applications have reached a level of sophistication where data are required from many national and international databases. A particular application may require simultaneous access to land use, climate and atmospheric data. These data could come from widely distributed holdings. Furthermore, other types of geophysical information could be needed to complement the application. The data sources are often domestic, but increasingly they include foreign data banks as well.

While Canada has extensive national and regional data holdings, many user studies have pointed out serious weaknesses in the data management infrastructure which hamper or preclude access to data. When data can be acquired, integration of the data is a major and costly task. To address these shortcomings in the Canadian earth observation data services, the CEONet program is being developed collaboratively by science based departments with major data holdings. The proposed CEONet program would improve access to data by providing an interoperable earth observation catalogue system for Canada, and by providing sample datasets and tools that will simplify catalogue access and visualization. It will also promote standards that will facilitate data integration.

Keywords: Data Management, catalogues, earth observation

Background

Canada is rich in national databases of interest to both domestic and international users (Noetix and Roger Buxton Associates, 1993). The Canadian archive of remotely-sensed imagery contains more than 1.2 million images covering the Canadian landmass and territorial waters many times since 1972. Canada is also the custodian for several international databases such as the World Data Center for Drifting Ocean Buoy Data. Although each of these data bases has value as a separate entity, their value, if combined and linked to multidisciplinary users, would be greatly enhanced.

Several independent user studies have identified serious weaknesses in the Canadian earth observation data management infrastructure that hamper or preclude access to the data (Noetix, 1993). When the data can be acquired, integration of the data is a major and costly task. Further, these studies have identified a common set of requirements: users need to find out rapidly and easily what data are available that may bear on the application of interest; they need to obtain the selected data and associated information products in the time frame and form that will make the data usable for the application; and they need to combine the data with those obtained from other sources, including holdings in national and international archives (Long Term Space Plan Interagency Working Group, 1992a). Typically, this step involves registration of the data in a common spatial framework and map projection. The integrated data set can then be analyzed to prepare new products or information sought by the user. Importantly, the user should be able to go through the above steps conveniently and effectively, employing tools available in his/her office.

Program Elements

The Canadian Earth Observation Network (CEONet) is being developed collaboratively by science based departments of the federal government to address these shortcomings in the Canadian earth observation data services. CEONet will bring together the remote sensing data archives and other databases, and will make their contents more accessible to users. It will also give Canadians access to other international databases, and will facilitate foreign access to Canadian data (Long Term Space Plan Interagency Working Group, 1992b).

The proposed CEONet program has the following elements:

Canadian Earth Observation Catalogue System: The objective of this element is to develop a network of Canadian databases that will support "one stop shopping," where a user can enter the network, locate all the relevant data in distributed databases, and obtain or place an order for the required data. This objective will be achieved by developing interoperability standards and software that will provide users with a standard access method to data in the major Canadian earth observation archives, and will provide interoperability with international systems. Figure 1 is a conceptual view of how the CEONet user communities will be linked.

Integrated Thematic Data Bases: The objective of this element is to promote the use of earth observation data to a very broad user community including: resource managers, scientists, educators and students. This will be accomplished by preparing thematic data packages containing baseline datasets that are important for many applications. These data will be converted to a standard form suitable for easy distribution, integration and processing with data from other sources.

Workstation Infrastructure Development: The objectives of this element are to provide the user with an integrated software environment to access and use earth observation data, and to provide industry with an advanced technology platform for product development for the earth observation market. This will be done by developing an integrated set of software tools for catalogue browse and visualization, for data ingest and format translation, and for data modelling and visualization.

Program Rationale

CEONet will address the data access problems identified in several recent studies and will contribute to the operational application of remote sensing and to the development of the technology for Earth observation. Turnaround time for finding, ordering and receiving data sets will be reduced, thus allowing the users to spend more time on their applications. With CEONet operational users and scientists will be able to go beyond their own internal data sets to gather the information they need through electronic networks, databases and distributed browsing systems. In addition the interface for new Earth observation missions will be in place, making data from all missions more quickly accessible.

Effective management of Canada's natural resources requires reliable information that can often be provided by earth observation. Canadian government and agencies have spent many billions of dollars collecting data to manage and monitor Canada's land and waters. CEONet will ensure that these data and the remote sensing data provided by the Canadian Space Program are used to their fullest potential.

CEONet will be a strategic source of new technologies for Canadian industry in a global knowledge-based economy. The value-added industry will use the CEONet infrastructure to develop and export new services for the management of earth resources and the environment. The ground station technology industry will incorporate the CEONet approaches and techniques in data delivery systems for the export market. This program is essential for Canadian industry to retain its leadership role in the application of technology derived from the Canadian Space Program and its share of the export markets.

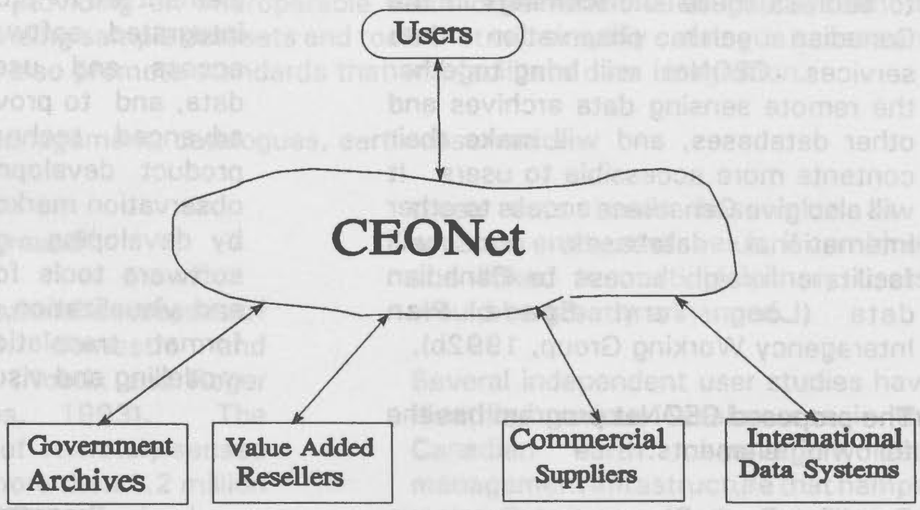


Figure 1

Program Status

Four agencies (the Canadian Space Agency, the Canada Centre for Remote Sensing, the Department of Fisheries and Oceans and the Atmospheric Environment Service) are funding a study to develop the requirements, conceptual design, implementation plan, costs, benefits and commercial opportunities of CEONet. The contract to complete this study was awarded to a group of Canadian companies led by

MacDonald Dettwiler and Associates of Richmond, B.C., in February 1994. This Study will be completed in October 1994. Development of subsequent phases of CEONet is dependent on the approval of this program in the Canadian Long Term Space Plan.

Conclusion

CEONet will increase and accelerate the use of Earth Observation data and the integration of space remote sensing in operational resource and environment management programs. At the same time Canadian industry will gain the ability to generate new knowledge-based products and services for the more globally-interdependent economy of the 1990s.

References

Noetix and Roger Buxton Associates, 1993, "Preparatory Study for CEOnet - Phase 1."

Noetix, 1993, " Preparatory Study for CEOnet - Phase 2."

Long Term Space Plan Interagency/Industrial Working Group 1992a, " Earth Observation Needs and Clients Subcommittee Report."

Long Term Space Plan Interagency/Industrial Working Group, 1992b, "Earth Observation Working Group Data Management Report."

Réseau canadien d'observation terrestre (CEONet)

Résumé

Les applications d'observation terrestre ont atteint un degré de complexité qui les oblige à tirer leur information d'un grand nombre de bases de données nationales et internationales. Une application particulière peut exiger un accès simultané à des données sur le climat, l'atmosphère et l'utilisation du terrain. Ces données peuvent venir de bases dont le contenu est distribué à grande échelle. De plus, d'autres types de renseignements géophysiques complémentaires peuvent être nécessaires. Les sources sont souvent canadiennes, mais elles comprennent également des banques de données étrangères.

Au Canada, même si on trouve un grand volume de données dans des bases nationales et régionales, de nombreuses études des utilisateurs ont révélé des lacunes importantes dans l'infrastructure de gestion des données qui gênent ou empêchent le transfert de l'information. Le programme proposé, CEONet, peut permettre d'améliorer l'accès à ces données des deux façons suivantes : premièrement, en fournissant au Canada un système de catalogage de l'observation terrestre offrant des possibilités d'interfonctionnement et, deuxièmement, en fournissant des outils et des ensembles de données d'essai qui simplifieront grandement l'accès et la visualisation et qui encourageront l'adoption de normes visant à faciliter l'intégration des données.

DEFINING THE ARCHITECTURAL DEVELOPMENT OF EOSDIS TO FACILITATE EXTENSION TO A WIDER DATA INFORMATION SYSTEM

Mark Elkington[§], Richard Meyer[¶] and Gail McConaughy[¥]

[§]*Earth Observation Science Ltd, Hughes Team ECS Project Office, 1616 McCormick Drive, Landover MD 20785*[*];

[¶]*D&M Associates, Hughes Team ECS Project Office, 1616 McCormick Drive, Landover MD 20785*[*];

[¥]*NASA Goddard Space Flight Center, ESDIS Project Office - Code 500, Greenbelt, Maryland*

Commission II, Working Group 2

KEYWORDS: Global Change, Information Systems, Earth Observation

ABSTRACT

To ensure that the Earth Observing System Data and Information System (EOSDIS) can play a role in earth science information systems that are likely to emerge in the next century, it is important that a suitable architectural direction is established from the outset of its development. The paper describes an open architectural concept under development by NASA for EOSDIS which supports site autonomy and independent, evolutionary development of components to improve the services offered to users. This concept is intended to ensure that EOSDIS' data sets and services can form part of a future international earth science system, but also offers several advantages for the future evolvability of the system itself.

1. INTRODUCTION

[1]NASA's Earth Observing System (EOS) is a long-term, multi-disciplinary research mission to study the processes leading to global change and to develop the capability to predict the future evolution of the Earth system on time scales of decades to centuries (Asrar and Dokken, 1993). The EOS Data Information System (EOSDIS) provides computing and network facilities to support the EOS research activities, including data interpretation and modeling; processing, distribution, and archiving of EOS data; and command and control of the spacecraft and instruments.

Although EOSDIS will eventually contain an enormous amount of valuable Earth science data, there are other sources of information that are essential to the study of climate change. Of critical importance are holdings of other Global Change agencies, such as NOAA, USGS, etc. and other international organizations. The Inter-Agency Working Group for Data Management for Global Change Data (IWGDMGC) are currently in the process of defining the Global Change Data and Information System (GCDIS) intended to provide linkages between data services through a common set of interoperability services. NASA is actively participating in these efforts.

In addition, there is also a growing interest by earth scientists in the possibility of developing information systems for earth science data which not only encompass the major data repositories but also enables users to take an active part in the information system, by providing data/services to the system (i.e. UserDIS). This approach seeks to encourage the scientific return from the investment in data and information systems by ensuring that the scientists are an integral part of the system.

Although NASA does not have the responsibility for developing either GCDIS or UserDIS it wants to make sure that its development of EOSDIS can support both of these evolutionary paths. This implies taking an architectural direction which opens EOSDIS so that it can be included within wider data systems and identifying architectural components which EOSDIS might contribute to these systems. The remainder of this paper discusses EOSDIS relative to GCDIS/UserDIS

This paper summarizes the results of NASA's preliminary architectural investigation, currently in progress. This paper presents high level user issues related to a generalized data and information system followed by an outline of an architectural concept for such a system. This is followed by a discussion of the major issues that would need to be resolved for the development of such a system.

2. ARCHITECTURAL ISSUES

The nature of Global Change and earth science research in general lead to some key architectural drivers for a data and information system to support this research:

- The organizations which participate in the network are autonomous entities, and the architecture should intrude upon autonomy to a minimal extent. For example, the architecture cannot dictate how organizations will manage their data, their networks, and their users internally.

* The work of the first two authors was funded as part of the Hughes Applied Information Systems contract to NASA for the EOSDIS Core System Project (contract NAS5-6000)

- Developing experiments, instrumentation, algorithms, and hence new kinds of data is an integral part of scientific research. Different science disciplines may have valid preferences for different or new data formats and tools. The architecture cannot make adherence to strict data interchange format standards or the use of specific tool sets a precondition of network operation.

- Scientists collaborate on research projects and exchange scientific information in many different ways. For example, scientists may want to obtain the latest set of re-calibrated satellite data or the outputs of a revised earth atmospheric model as inputs to their own analysis. The architecture should facilitate this kind of collaboration and exchange and extend to new ways of collaboration which future technologies may enable.

- A key objective of the Global Change Research Program is the establishment and maintenance of a high quality set of earth science parameters for an extended period of time which have research community consensus [from 'Global Change Data and Information System (GCDIS): A Draft Tri-Agency Implementation Plan', DOI/NASA/NOAA, March 1992]. The research community as a whole has been challenged to cooperate in the validation, upgrade, and description of this data. The architecture needs to support the research community in finding relevant data, in the analysis and critical review of this data and the publishing and dissemination of new and revised data products.

- The science expertise is distributed on a world-wide scale. The architecture should make it possible to distribute the appropriate functions and data to where the expertise resides. The cooperative endeavor envisioned by the GCRP also means that the architecture must allow researchers to take advantage of the distributed functions and data in collaborative efforts.

- Scientists would like to access scientific information in many different ways. The architecture should extend to new ways of data access which future technologies may enable. Data volume makes it impossible today to perform large scale searches on science data based on their contents, but the rapid evolution of processing and storage technology may change that in the not too distant future.

- The collaboration between geographically distributed researchers is limited by current communications facilities to file and mail exchanges. In the future, it may be possible to exchange data in real time, and view and browse them in a coordinated fashion while communicating annotations and comments. Such developments will have a profound impact on how earth science information is accessed and used for research.

- It is a characteristic of global change research that it tries to correlate information which spans long periods of time, and that experiments may go on for many years. Therefore, the architecture must be able to accommodate the long term evolution of technology and its application to science, but at the same time it must provide some measure of stability, e.g., in terms of backward compatibility.

The system architecture for EOSDIS must support both users and service providers. The concept of a service is synonymous with the concept of data, since data can only be accessed through a service (e.g. ftp, DBMS, etc.). The users of the system want to be able to find and access relevant services within the system as efficiently as possible. Service providers want to be able to support the mission objectives in terms of capturing and maintaining the important data sets, and ultimately providing the best services possible to the user community. To achieve the latter goal each provider needs the flexibility to organize their data and services in the most appropriate way for their user community and be free to re- configure existing services and add new ones to accommodate new user requirements and/or new technological capabilities.

In this scenario of autonomous service providers, there will inevitably be the potential for incompatibilities between the services available from the system and a user's query, or the tool they are using to access the service(s). The system must recognize that incompatibilities will exist and assist the user in overcoming them as effectively as possible.

An important focus for EOS is interdisciplinary science. This will lead to user requests which cannot be resolved by a single service or even a single service provider. The system architecture must therefore support the concept of multi-site requests, which must be partitioned and managed between several services. An example of this type of request and how it might be managed is shown in Figure 2.

The main objectives of the architecture are, therefore, the definition of (a) capabilities which let a scientist locate, obtain, or use resources which are available in the network (e.g., tools and data); (b) features which would help a scientist cope with the ensuing problems, e.g., of differences in data formats, terminology, and tool input and output requirements; and (c) support functions which would make it easier for scientists to collaborate on research projects across the network.

In widening the constraints of the architectural concern from the earth observation focus of EOSDIS towards global change research, to the wider earth science focus of GCDIS and UserDIS, there are many issues which are wider in scope than if the EOSDIS requirements alone were being considered.. These are discussed in the GCDIS/USerDIS study (NASA, 1994a) and are summarized below:

- There will be considerable variety in the user objectives, missions and priorities. Users will also be data providers.

- The architecture must not assume a common information model or management system. The adoption of evolving standards should be encouraged within the GCDIS / UserDIS community, though this should be achieved through participation in the standards process by the agencies and organizations involved rather than the development of specific standards for these systems.

- There should not be any restrictions on the number of providers, their location and the data/services they provide. The system must be able to cope successfully with dynamic data and network topology.

- All responsibilities for system management or development policies and authorities will be voluntary, though within a part of the system, such as EOSDIS, can be mandated by some management authority. The architecture must therefore accommodate autonomously managed provider sites and not assume a single management approach to development, operation, user authentication or data protection. In particular the system should not depend on the availability of network wide management information.

- The data management solutions should be scalable, and cost effective to scale. The design of components should avoid limits on capacity which preclude low-end providers or restrict what high-end providers can offer.

- The architecture must help the user work effectively within an environment characterized by variability of quality in terms of responsiveness, reliability, accuracy, availability, and throughput.

Taken together, these characteristics present some significant challenges to the design and development of EOSDIS if it is to be part of a wider data system and be a major supplier of components for such a system. It is important therefore that EOSDIS establishes what it is able to achieve within its cost and schedule budgets, and leaves open to future development those aspects beyond its scope.

3. ARCHITECTURAL CONCEPT

The architectural concept for EOSDIS is shown in Figure 1 and described in further detail in NASA 1994b. It can be divided into three layers: the client layer, the service provider layer and the interoperability infrastructure. Individual sites, which may host one or more of these layers, are heterogeneous and autonomously managed. The user layer is characterized by client environments, which may be interactive (e.g., workstation graphical interface) or process environments (e.g., analysis algorithm). The interoperability layer is characterized by a set of distributed services which assess user needs against service offerings and connect the user with appropriate service providers. Finally, the provider layer is characterized by organizations who choose, or are mandated by their management authority, to provide a set of services related to data collections or to computer resources that they can offer. This includes the traditional data center concept and also specialist value-added service providers, whether commercial or government related (e.g. education specialists). Since the service provider layer must allow autonomous management and development, the details given here are limited to those which allow sites to interoperate. The architectural concept then, is in essence the interoperability infrastructure (the Intersite Architecture) and how the user and data provider services interface to this infrastructure.

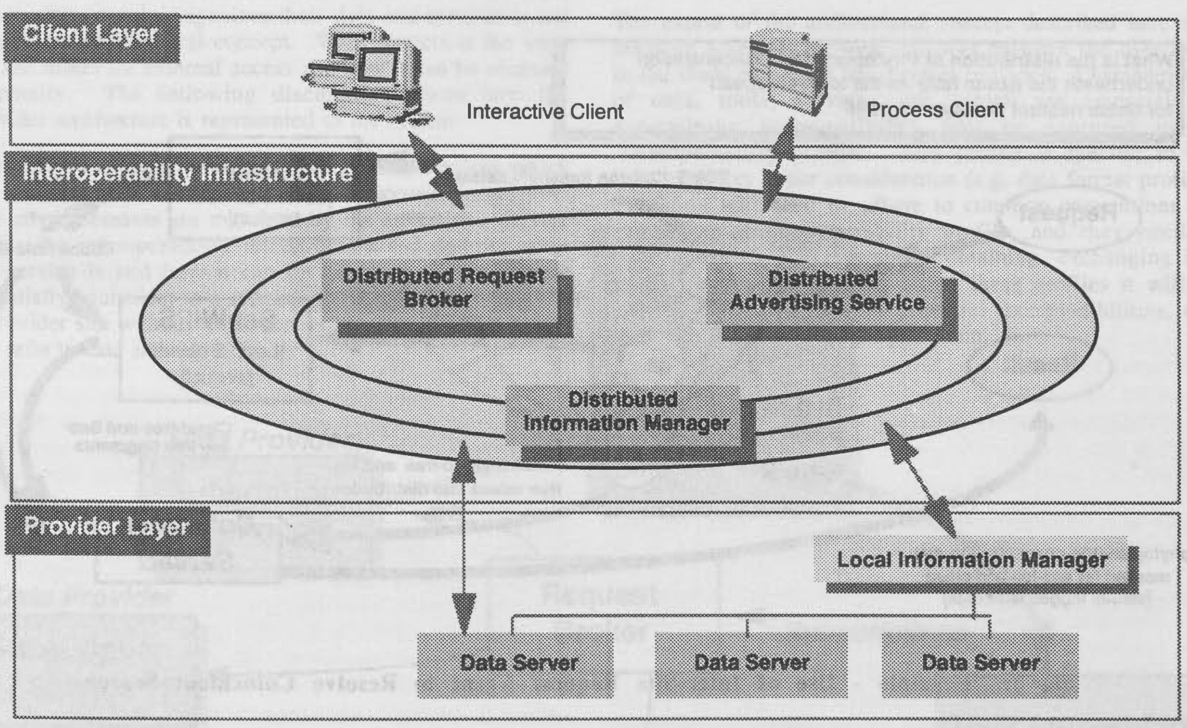


Fig 1-: Conceptual Architecture Overview

3.1 Intersite Architecture

The intersite architecture is necessary to connect the services offered by providers to needs expressed by users in terms of requests to the system. Three classes of software component are envisaged for this layer:

Advertising Service: The services offered by providers on the network need to be advertised to users. The advertising service is used by other components in the intersite architecture to perform their function.

Request Broker Service: This service performs the matching between a user request and the services offered. In the many cases the user (e.g., a process) may have specified the service to be connected to, in others the broker will have to parse a 'request description' and use the advertising service to establish which services could satisfy the request. The request broking activity might involve user interaction.

Distributed Information Manager: Where multiple sites are needed to resolve a request then an Inter-site Search Service is required to manage the process. The service will break up the query if necessary and generate a plan containing sub-requests which will be processed at individual sites. The sub-requests will generally be characterized by queries and operations, where an operation is usually some manipulation of results to provide output in the form or context requested by the user. The optimization of the division of a user request into a request plan is a difficult problem and in the short-term might involve the user in the planning process.

An example of the use of the Distributed Information Manager to handle a coincident, content-based search is shown in Figure 2. In this case, the request is divided into three sub-requests: a query is made on the TOMS binned product to extract an ozone mask of the ozone hole for the required date, the mask is transferred to the second service provider which extracts cloud and sea-ice free fragments of the binned phytoplankton concentration product derived from Sea-WiFS data, which are then mosaiced on a separate compute server to form a single multi-date concentration map.

Clearly this is not a trivial problem, and represents the vision of what should be possible in the future rather than what can immediately be developed. Issues such as optimization of the query plan, incompatibility of vocabularies, etc., all need to be addressed before such a vision is achievable. However there are clear evolutionary steps on the way to the vision, each of which give the user more support for resolving science questions. The purpose of the intersite architecture is to enable this evolution without a need to leave the architecture framework.

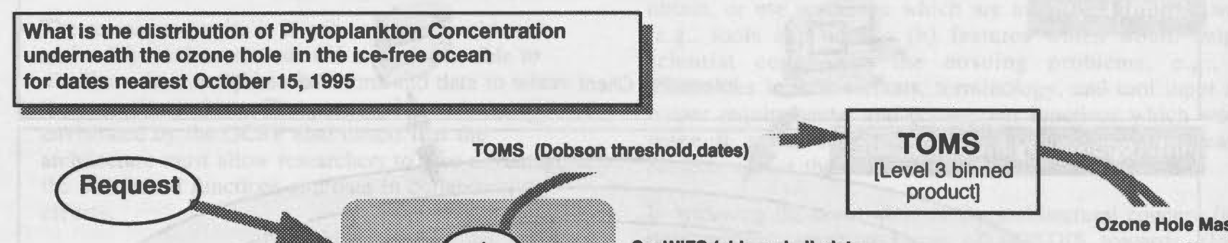


Fig 2. Example - Use of Inter-Site Request Agent to Resolve Coincident Search

The main issues related to the development of these three intersite components are related to vocabulary management and mapping, service and client incompatibility management and support for multi-protocol access to a dataset. These issues are considered further in NASA, 1994a.

The interoperability layer assumes that each of its services can be distributed. There will probably be more than one example of a request broker or distributed information manager. They can differ in the scope of their capabilities or to which services and providers they give access; that is, intersite services can themselves be heterogeneous. For example, a distributed information manager might provide relatively basic inter-site request capability across a wide range of services, or more sophisticated capabilities over a smaller subset.

3.2 Provider Interfaces to the Intersite Architecture

Each provider site chooses, or is mandated by its management, to provide certain services to the system community, or some part of that community. In many cases a service would be related to data set(s) held at the providers site, but it is not mandatory; a service potentially could access data held at another site, or even provide a service for data passed to it from another site as part of a request.

The main issues to be resolved in the interfacing of a provider services to the entire system are briefly described in this section; they are:

- autonomous internal organization
- advertising services to users
- support for searching
- support for notifications to users when new data and/or services are available - subscription services
- support for incompatibility management
- the role of data servers.

Provider sites should be allowed to *autonomously organize* and manage their internal services and data to permit political and technical flexibility and therefore a definition of how the provider organizes their data and services is not part of the architectural concept. What matters is the what the site offers for external access and how it can be accessed externally. The following discussion defines how the provider architecture is represented to the system.

Sites *advertise* their services to the system; services which are not advertised do not exist from the system perspective. The advertisements are managed by the advertising service within the interoperability infrastructure and describe what the service is and how it can be accessed. Since data is essentially equivalent to a service (a user can't access data at a provider site without some sort of service), advertisements can refer to data and services.

Sites which offer *data searching* must have an external interface for accepting searches (e.g. from the Distributed Information Manager), and a service for processing these requests. This is called the Local Information Manager, and is equivalent to the Distributed Information Manager, in that it resolves inexact search requests into exact queries which can be placed on individual data servers. This might involve some interaction with the user.

Sites may let users (or programs) subscribe to data which they store or distribute. This might be used, for example, to inform a user when new data is acquired for their particular study area which exceeds a certain quality threshold. To make this available the site must offer an external interface for a *subscription service*. Users can indicate their areas of interest in a search language which are then routed like other searches to services which can monitor the subscription. These subscription services then send notifications of new information items in those areas matching the user's request whenever they are found as part of the routine processing and archiving activity (see Figure 3).

The importance of this concept is that it will free scientists from mundane information/data hunting, allowing them to specify interests periodically and then receive notifications of new data and/or services that are relevant to those interests. The acceptance of this functionality and its implementation could have a major impact on the way data and information systems are developed for the future.

The nature of the architectural concept described here will result in potential *incompatibilities* between the user tools in the client layer and the services provided. Characteristics of data, tools, services, etc. which are essential for determining incompatibility will be captured in an 'interoperability profile'. This profile is specific to the type of object under consideration (e.g. data format profile). Providers will need to adhere to common conventions for describing an interoperability profile, and they need to provide external interfaces for obtaining, exchanging and negotiating such profiles. Using these profiles it will be possible to warn users of potential incompatibilities, and offer advice on how to mitigate them.

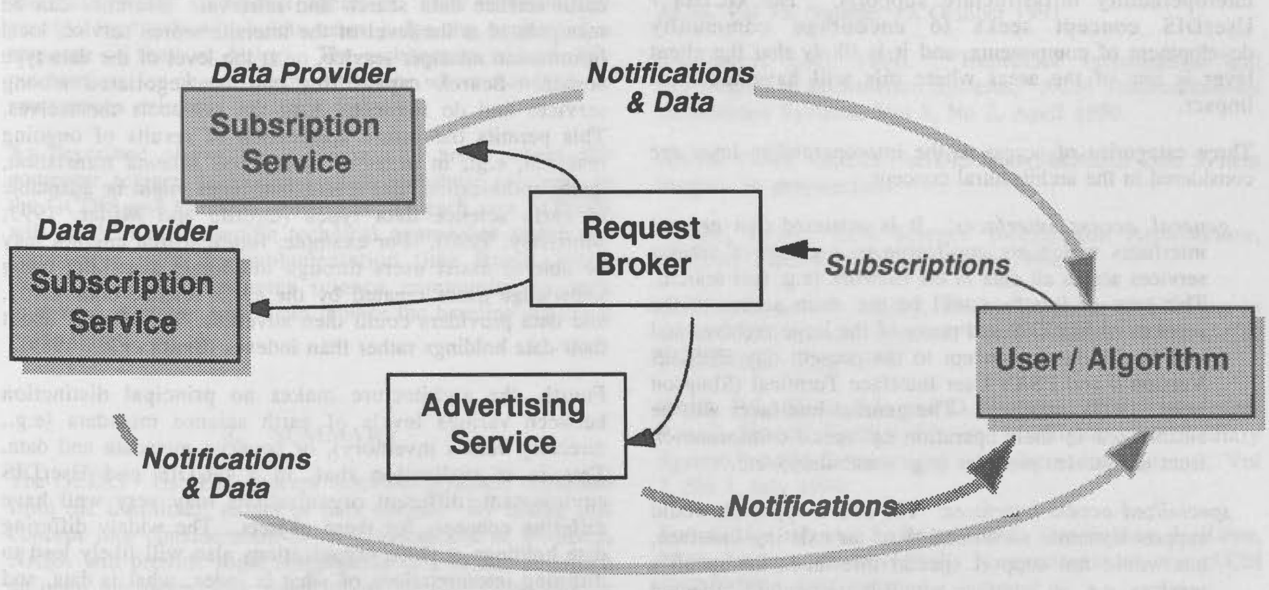


Fig 3: Subscription Service Concept

At each provider site the messages passed from the client layer to the interoperability structure need to be interpreted and acted upon. This is achieved through two routes. The most direct is a message directed at a *data server* which has knowledge of all the services related to its data set and passes the message to the correct service. At sites providing several data sets, the message might need to be interpreted against several data set servers and this would be handled by the Local Search Service. This is equivalent to the Distributed Information Manager, but only deals with the data held locally.

Data at a site is organized into one or more collections of related items. Each collection forms a data set which will contain both data and meta-data, the discrimination between these being provider and data set specific. To each data set one or more services are attached; the services may operate on all types of data in the data set or only one part of it, e.g., a relational database management system (DBMS) to an inventory table. Descriptions of these services and the parts of the data set they operate on are passed to the advertising service, establishing a 'data scope' against which users requests can be evaluated against. These services are called 'type' services in the concept since one type of service may be related to all data of that type (e.g., a text query service could be used for all text data within one data set and across all data sets at one site). Careful design of the type services should mean that they are adaptable by other providers for similar data. It is possible that a particular type of data would have more than one 'type' service associated with it, (e.g., two different text query services to support different access protocols) and that a type service would deal with more than just a single data type (e.g., an OO DBMS could deal with inventory data and the associated browse data).

3.3 User Interfaces to the Services

As for the provider layer, the architectural concept does not mandate how the user interfaces should work, only that they are compatible with one or more of the protocols that the interoperability infrastructure supports. The GCDIS / UserDIS concept seeks to encourage community development of components, and it is likely that the client layer is one of the areas where this will have the most impact.

Three categories of access to the interoperability layer are considered in the architectural concept:

general access interfaces: It is assumed that general interfaces which are applicable to a group of similar services across all data in the network (e.g. text search). This type of interface will be the main access to the services of EOSDIS and many of the large archives and will be similar in concept to the present day EOSDIS Version 0 and ESA's User Interface Terminal (Simpson *et al.*, 1993) interfaces. The general interfaces will be customized in their operation by specific information from the service provider (e.g. vocabulary), etc.

specialized access interfaces: The above approach would support dynamic modification of an existing interface, but would not support special interfaces for specific services, e.g. an interface which is particularly oriented towards the coincident location and analysis of sea surface temperature 'images' and sub-surface profile measurements of temperature and salinity. In this case the service provider might be able to provide a software module which would be dynamically linked into the user's interface and provide a completely specialized interface, configured to a specific service/provider.

object access interfaces: Finally, objects resulting from previous queries should be capable of initiating further service requests. For example a search of an image inventory might result in a results object which contains the inventory records matching the query and a reference in the object which would enable a user to automatically initiate a browse service and review the image being referenced by one of the inventory records.

4. DEVELOPMENT ISSUES/APPROACH

The EOSDIS architectural concept described here offers several important advantages for some future development of GCDIS and/or UserDIS. First, data providers have complete freedom of choice as to how they wish to organize data into types. They may want to use EOSDIS provided type services or data type servers, or they may decide to create their own and link them into the local search and retrieval service software. They may even decide to replace all data types with a fully integrated database management system (and appropriate schema); several research projects and commercial ventures in the area of object databases are in progress and may mature during the lifetime of the global change program (Stonebraker and Rowe, 1986; Haas *et al.*, 1990; Lamb *et al.*, 1991).

Secondly, the concept supports the inclusion of legacy systems into the network. A site is only required to advertise only those services it wishes to support in the GCDIS/UserDIS context; there is no minimum set of services which a site must provide. For example, a site may only be able to offer text search and file transfer, but would still be able to contribute to the network.

Thirdly, the architecture provides an open ended approach to earth science data search and retrieval. Searches can be manipulated at the level of the intersite search service, local information manager service, or at the level of the data type server. Search capabilities can be negotiated among services and do not enter into the protocols themselves. This permits the future application of results of ongoing research, e.g., in areas of query and schema translation, query optimization, and search languages might be adaptable to earth science data types (Ordille and Miller, 1993; Morrissey, 1990). For example, future search engines may be able to assist users through intelligent searching using knowledge bases created by the earth science community, and data providers could then advertise "knowledge" about their data holdings rather than indexes (Smith *et al.*, 1989).

Fourth, the architecture makes no principal distinction between various levels of earth science metadata (e.g., directory versus inventory), or between metadata and data. This is in realization that, in a GCDIS and UserDIS environment, different organizations may very well have differing concepts for these objects. The widely differing data holdings at these organizations also will likely lead to differing interpretations of what is index, what is data, and what can be searched within a reasonable time. Despite this variability in data it is important that searches on multiple data sets provide results which can be compared effectively by the user, and thus the query process will include mechanisms to ensure that the user receives the required form of result. The architectural approach shown facilitates the introduction of more powerful search strategies in the future (Hellerstein and Stonebraker, 1993; Haas, 1989).

Finally, the concept described above will encourage evolutionary and independent development of system components. By adopting a fully distributed architecture for all components and not mandating the details of the client interface and service implementations, the entire user and development community can participate in the development of components in each of the three layers. For example, computer science research may lead to the development of an improved intersite search agent. Users can then choose whether the new agent provides a 'better' service. If it does then, over time, it will make other agents obsolete. Moreover by establishing a conceptual framework which can accommodate the variability of the earth science discipline which can guide rather than constrain development of components, hopefully minimizing the 'not invented here' syndrome, it will encourage the development of components and support utilities (e.g. APIs) by the entire community.

Although the architectural concept seeks to strike a proper balance between the users' demand for decentralized capabilities and autonomy on one side, and complete anarchy on the other, a network of the type proposed for GCDIS / UserDIS poses significant issues in several system quality areas. For example, the accuracy of search results suffers as incompatibilities among the vocabularies and terms employed by different data providers increases. In an unmanaged network, there can be no expectations regarding service reliability, availability and response time. For example, some sites may respond to a search within seconds or minutes, others may not respond for days because the data provider experiences hardware problems.

The solutions to these types of problems are outside the scope of an architecture. They depend on the cooperation of service providers which, in a network like UserDIS, is voluntary. However, the architecture can include measures to facilitate the solutions. For example, EOSDIS will not make a reliable network, in which all sites are always available, a precondition for successful operation. The services will provide feedback which lets users judge the quality of a response (if they so desire). The architecture will provide mechanisms for characterizing situations where standards or conventions exist and are being followed.

As described above there are several areas where the computer science community could contribute solutions to the GCDIS and UserDIS challenges. In each area EOSDIS will need to pick specific technical approaches which are compatible with its implementation time frame, while encouraging the computer science community to seek improved solutions which can replace the baseline approach in the future.

5. SUMMARY

The GCDIS / UserDIS concept describes a radical departure from the traditional model of data system. By taking this concept into consideration in its development of EOSDIS, NASA will provide some components of a system in which an open interoperability standard can be used to acquire or provide data and services, enabling an information system to be developed that will operate more as a marketplace with positive competition than as a monolithic, monopoly that focuses on production and storage of data.

Such an information system should encourage evolutionary and independent development within a single framework on an inter-agency and international scale. Indeed its success depends on this complementary development. It should also provide more flexibility for accommodation of new user needs and taking advantage of emerging technological developments. Finally, it provides more flexibility to respond to the inevitable change in distribution, prioritization and funding policies over such a long-term undertaking as an earth science information system.

ACKNOWLEDGMENTS

This work was undertaken as part of Hughes Applied Information Systems contract to NASA for the EOSDIS Core System Project (contract NAS5-6000). Important recommendations which led to the approach described in the paper were received from the National Research Council's review of the EOSDIS plans.

REFERENCES

Asrar, G. and D. J., Dokken, 1993. EOS Reference Handbook, NASA, March 1993.

Haas, L.M., Freytag, J.C., Lohman, G.M., and P. Pirahesh, 1989. "Extensible Query Processing in Starburst," Proceedings of the ACM SIGMOD'89, June 1989.

Haas, L.M. *et al.*, 1990. "Starburst Mid Flight: As the Dust Clears," IEEE Transactions on Knowledge and Data Engineering, March 1990.

Hellerstein, J.M. and M. Stonebraker, 1993. "Predicate Migration: Optimizing Queries with Expensive Predicates", Proceedings of the ACM SIGMOD '93, June 1993.

Lamb, C., Landis G., Orenstein J., and D. Weinreb, 1991. "The ObjectStore database system," Communications of the ACM, Vol 34, Number 10, October 1991.

Morrissey, J.M., 1990. "Imprecise Information and Uncertainty in Information Systems," ACM Transactions on Information Systems, Vol 8, No 2, April 1990.

NASA, 1994a. GCDIS/UserDIS Study, EOSDIS Core System Project, in preparation.

NASA, 1994b. ECS Science Information Architecture, Working Paper FB9401V2, March 1994.

Ordille, J., and B. Miller, 1993. "Database Challenges in Global Information Systems," Proceeedings of the ACM SIGMOD'93, June 1993.

Smith, P.J., Shute, S.J., and D. Galdes, 1989. "Knowledge-Based Search Tactics for an Intelligent Intermediary System," ACM Transactions on Information Systems, Vol 7, No 3, July 1989.

Stonebraker, M. and L. Rowe, 1986; Stonebraker, L. Rowe, "The Design of Postgres", Proceedings of the ACM SIGMOD'86, June 1986.

LE DÉVELOPPEMENT DE L'ARCHITECTURE DE EODIS AFIN DE FACILITER LA TRANSITION VERS UN SYSTÈME D'INFORMATION DE DONNÉES GLOBAL

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif.

PROGRAMMING WITH CONSTRAINTS IN AN OBJECT-ORIENTED GIS

FRANCOIS BOUILLE

L.I.S.T., Boîte 1000, Université P.M.Curie, 4 Place Jussieu
75252 PARIS cédex 05, FRANCE

ABSTRACT

All geographical phenomenons may be easily structured in an object-oriented GIS. The HBDS model is one of the possibilities for doing that. Based upon six basical persistent abstract data types and their extensions, it allows to build a class-based integrated platform as a kernel for GIS design and development. It includes an expert system managing fuzzy rules and facts, and a neural engine allowing automated learning; but for expressing most of constraints in a GIS, the rules are not the universal tool, and the concept of "constraint" must be introduced as a new component which can be defined, handled and applied. It requires a special architecture and a specific vocabulary. This new tool must be used in cooperation with the others, such as the classical "condition" carried by any component, such as more complex tools i.e. the "rule" , the "process" and the "neuron", all concepts making a GIS fully dynamical. These constraints are particularly usefull when simultaneously dealing with several types; they may be strict or fuzzy; we may distinguish too constraints relevant to conditions which must be absolutely verified, and that relevant to conditions which must never happen. As all the components of the HBDS model, constraints have an intrinsic graphical representation, are organized into classes, present their own properties and relationships with other components and between contraints themselves. This new tool introduced in GIS provides a new way of expressing the constraints but requires to take care in programming..

RESUME

Tout phénomène géographique est facilement structuré dans un SIG orienté objet. Le modèle HBDS représente l'une de ces possibilités. Fondé sur six types abstraits de données persistants et sur leurs extensions, il permet de construire une plateforme intégrée orientée objet, constituant un noyau pour la conception et le développement de SIG. Il comprend un système expert gèrant des faits et règles flous, ainsi qu'un moteur neuronal permettant un apprentissage automatique; mais pour exprimer la plupart des contraintes dans un SIG, les règles ne constituent pas l'outil universel, et le concept de "contrainte" doit être introduit comme un nouveau composant qui peut être défini, manipulé et appliqué. Ceci demande une architecture spéciale et un vocabulaire spécifique. Ce nouvel outil doit être utilisé en coopération avec les autres, tels que la classique "condition" qui peut être portée par n'importe quel composant, tels que des outils plus complexes comme la "règle", le "processus" et le "neurone", tous ces composants donnant au SIG son aspect dynamique. Ces contraintes sont particulièrement utiles lorsque traitant simultanément de plusieurs types; elles peuvent être strictes ou floues,; l'on peut distinguer aussi les contraintes relevant de conditions qui doivent être absolument vérifiées et celles qui traitent de circonstances qui ne doivent jamais arriver. Comme tous les composants du modèle HBDS, les contraintes ont une représentation graphique intrinsèque, sont organisées en classes, presentent leurs propres propriétés et relations, soit avec les autres composants, soit entre contraintes elles mêmes.

KEY WORDS: constraint programming, knowledge base, neural model, object-orientation.

1. INTRODUCTION

Most of components in a GIS, for instance dealing with urban networks, are easily structured, captured, stored and processed in an object-oriented environment. Such a methodology allows to express most of properties and relationships between the components better than a relational approach could do. Nevertheless, most of o.-o. models have some difficulties in expressing the constraints between the geographical elements of a structure. Though the use of an expert system is an interesting way, rules do not provide a well-suited methodology for expressing and applying constraints on a geographical model. That is why we propose to introduce in the GIS area a new way of programming by using constraints.

2. BRIEF RECALL OF THE HBDS MODEL

HBDS means "Hypergraph-Based Data Structure; early designed in year 1977 (Bouillé 77a, 77b, 79), it was progressively developped for providing all the interconnected tools a GIS user may require. We may distinguish three things: -the model which is computer-independent, -the complete system, -and the applications, with two major fields: -the oil exploration and production, -the Geographic

Information Systems. In this last area, HBDS is the standard for geographical data structuring at the Institut Géographique National in France (Dassonville, 1991).

Basicaly, HBDS is composed of six basical abstract data types (indicated here as "a.d.t."), respectively named:

- –class, attribute of class, link between classes, these three a.d.t. composing the "structure skeleton",
- –object, attribute of object, link between objects, these three a.d.t. composing the "realizations".

Figure 1 shows the graphical representation associated to each a.d.t.

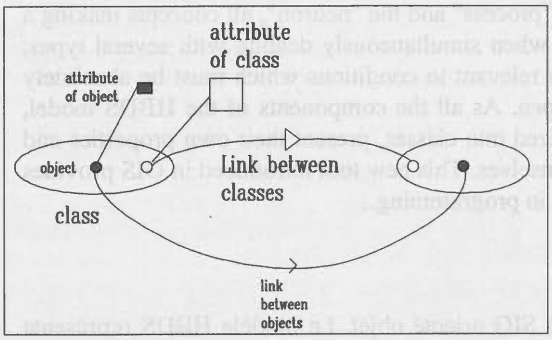


Fig. 1 : Graphical representation of the six basical abstract data types.

An attribute has a kind, a type, possibly predefined values, a possible unit when digital, a given type of fuzziness, and various other features we cannot detail in these few pages. The different kinds are: –scalar, –array (distinguishing vector, matrix and tensor of any size), –list (circular or linear) –compound (looking like a C–structure or a Pascal record).
Arrays may be built as "sparse", when there are many missing values, for instance, and this capability concerns vectors, matrixes and tensors.
The possible types are : –integer, –float, –complex, –gauss (Gauss complex), –quaternion (=hypercomplex, very usefull in robotics), –rational, –boolean, –character, –string (of any length, extensible as necessary without previous declared length).
Digital types may be built as "long", thus having any number of digits.
Links between classes carry a potential relationship, whereas links between objects represent an effective relationship verified by a couple of objects.

In HBDS, all the a.d.t. are considered as implicitely persistent.The first extensions are three a.d.t., respectively named : hyperclass (a set of classes), hyperattribute (attribute carried by a hyperclass and inherited by all the classes included in this hyperclass), hyperlink (link starting from a hyperclass, or arriving on a hyperclass, and thus inherited). Figure 2 shows the graphical representation of these new a.d.t.

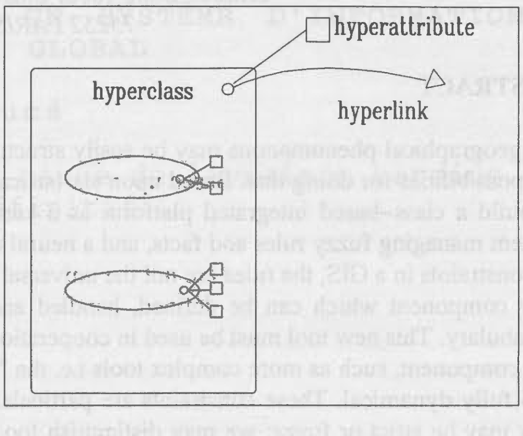


Fig. 2 : The three first extensions.

Hyperclasses may intersect or embed others, thus providing a multiple inheritancy. The second extensions are respectively named embryo, prototype and structure. An embryo looks like a basical a.d.t. of the skeleton, but with no realization; it is used for assembling with others, composing a prototype (Bouillé, 1993b); it is just a model, which copies will provide structures, each structure receiving afterwards more specific attributes and/or links corresponding to its specific phenomenon features. The figure 3 shows the graphical representation of a prototype, named "GRAPH", representing all the topological relationships between vertices (improperly named "nodes"), arcs (also sometimes named segments) and areas, as it may be found in most of geographical themes.

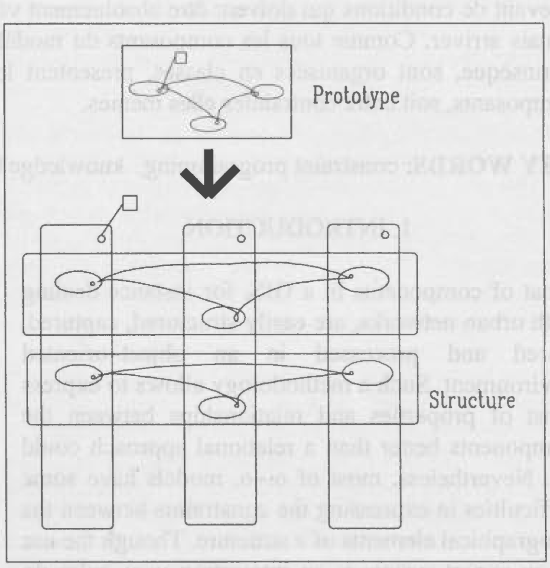


Fig. 3 : From prototype to structures.

The third set of extensions consist of building new components, corresponding to classes of rules, classes of processes and classes of neurons.

Rules are considered as objects of classes which have been previously included into a hyperclass named "RULE"; thus, they inherit of the attributes and links of this hyperclass, namely: –premis (boolean expression, or fuzzy expression), –consequent if yes (any set of executable statements), –consequent in no (likewise), –a preparing part named "PROLOG" (exec.statements), –a closing part named "EPILOG" (likewise), –an explanation which is purely textual, –various coefficients such as: –askability, –editability, –priority, –modifiability, –lauchability, –priority, –reliability, etc... Moreover, rules present relationships such as "actives" and "inhibits", because a rule may active another one which was not *a priori* activable, or on the contrary, an activated rule may inhib another rule which is launchable but not compatible with the first one. The figure 4 shows the graphical representation of the hyperclass RULE.

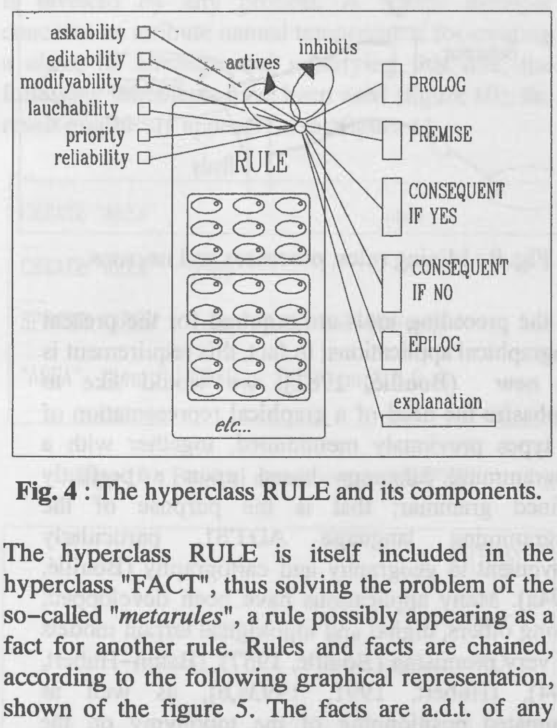


Fig. 4 : The hyperclass RULE and its components.

The hyperclass RULE is itself included in the hyperclass "FACT", thus solving the problem of the so-called "*metarules*", a rule possibly appearing as a fact for another rule. Rules and facts are chained, according to the following graphical representation, shown of the figure 5. The facts are a.d.t. of any type, such as attributes of classes, links between classes, classes themselves, as shown on the figure 5; they may be too objects or anything else. The expert system included in the HBDS kernel is not composed of a single inference engine, but composed of several ones which run forward and backward. These engines accept fuzzy facts and fuzzy rules. The engines are themselves included in the knowledge structure. The whole expert system is designed without the obsolete concepts of backtracking and pattern-matching (Bouillé, 1984a,b, 1988, 1991b).

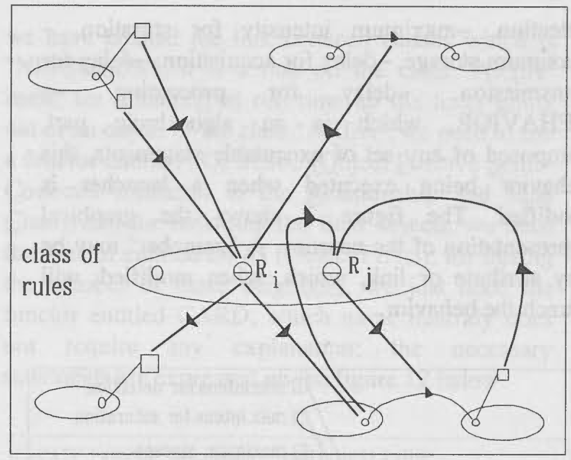


Fig. 5 : Graphical representation of the interaction of rules and facts.

Discrete processes are considered as objects of classes which have been previously included into a hyperclass named "PROCESS"; thus they inherit of all the attributes of this hyperclass, namely : –a clock, dealing with the time at the scale of a possible simulation (nanoseconds as well as centuries or b.y.), –state (which may be active, passive, hold, terminated, –body, which is the algorithmic part comosed of any set of executable statements, working on a *coroutine* mode, –a local sequence counter, according to the term firstly introduced in the SIMULA 67 programming language. The figure 6 shows the structure of the Hyperclass "PROCESS".

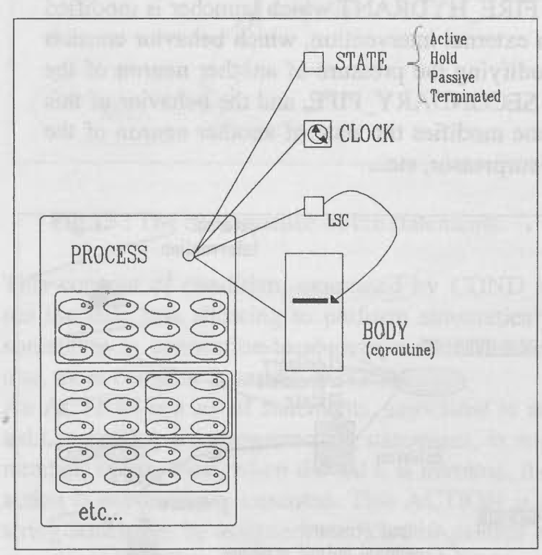


Fig.6 : The hyperclass PROCESS and its components.

Neurons too are considered as objects of classes which have been previously included into a hyperclass named "NEURON"; thus they inherit of the following attributes: –minimum intensity for detection, –maximum intensity for sturation, –maximum storage, –delay for acquisition, –delay for transmission, –delay for processing, –a BEHAVIOR, which is an algorithmic part, composed of any set of executable statements, this behavior being executed when a launcher is modified. The figure 7 shows the graphical representation of the neurons. A "launcher" may be any attribute or link, which, when modified, will launch the behavior.

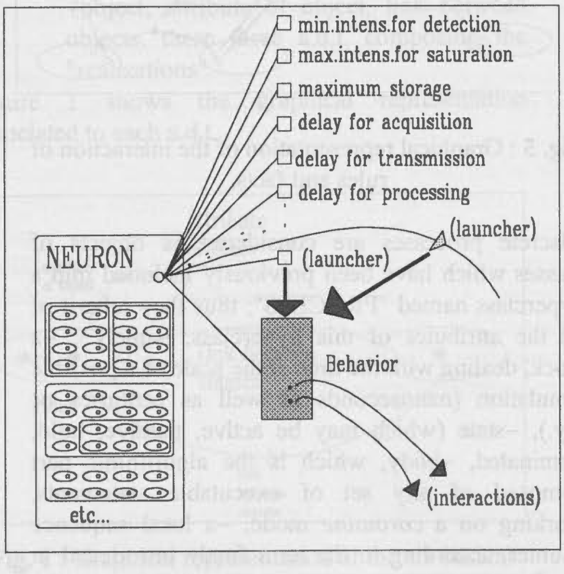


Fig.7 : Hyperclass NEURON and its components.

For instance, on the figure 8, we see a neuron of the class FIRE_HYDRANT which launcher is modified by an external intervention, which behavior consists of modifying the pressure of another neuron of the class SECONDARY_PIPE, and the behavior of this last one modifies the state of another neuron of the class surpressor, etc...

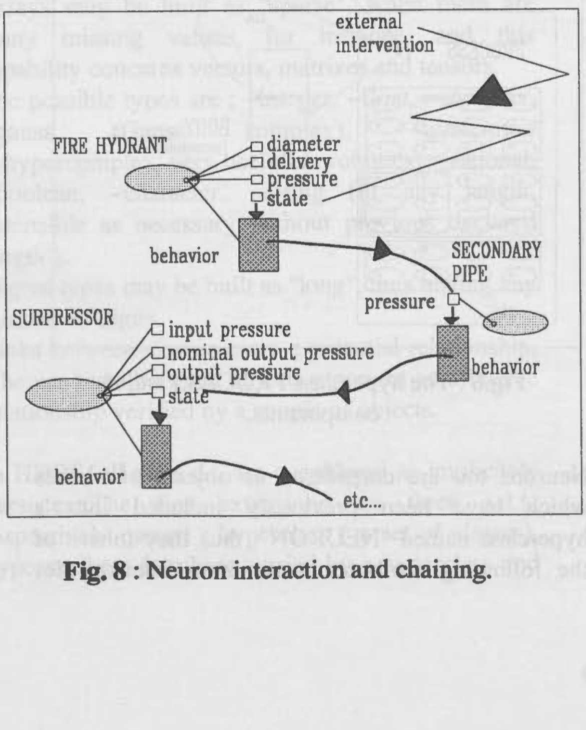


Fig. 8 : Neuron interaction and chaining.

Neurons too may be fuzzy (Bouillé, 1991b); they are very usefull in terrain modelling (Bouillé, 1992). Rules, processes and neurons have not exactly the same goals; though anyone may sometimes replace another one (Bouillé, 1993a, 1993c), they are more dedicated to specific topics, respectively expert system, simulation, automated learning. Nevertheless, they may interfer, a rule possibly modifying a fact which is a launcher for a neuron, the behavior of this one then activating a process which body will modify another a.d.t., as represented on the figure 9.

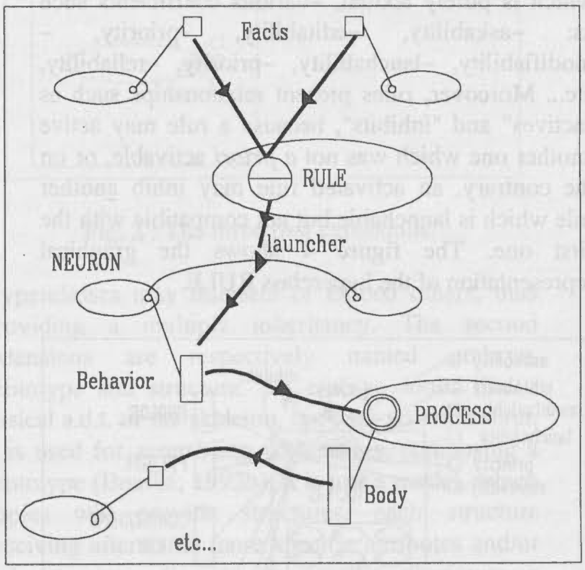


Fig. 9 : Mixing rules, processes and neurons.

All the preceding tools are required for the present geographical applications; in fact, this requirement is not new (Bouillé, 1981); we would like to emphasize the need of a graphical representation of all types previously mentionned, together with a programming language based upon a perfectly defined grammar; that is the purpose of the programming language ADT'81, particularly convenient to geography and cartography (Bouillé, 1994a). Many applications have been developped; among others, digital and topological terrain models are very promising (Bouillé, 1987), (Baton–Hubert, 1994), (Hubert, 1991, 1993a,b), as well as automated positionning of the toponymy on the maps (Titeux, 1989).

The complete kernel of the HBDS system includes several tools interconnected, we just here mention:
–a multi–level indexed sequential file system,
–a very large object–oriented database,based upon some decisive criteria (Bouillé, 1991a),
–a multi–engine fuzzy expert system managing an illimited set of persistent rules and facts,the database thus becoming a very large knowledge base,
–a simulation system with an engine managing an

illimited set of discrete processes, in connection to a second engine managing continous processes,
-a neural engine managing an illimited set of object-oriented neurons, possibly fuzzy, ensuring an automated learning, in cooperation with the expert system,
-a 3D stereoscopical animated graphical engine, in a multimedia context,
-a specific layer devoted to the robotics (Bouillé, 1986), in prevision of the large application of robotics to geography and cartography in the next decade.
The complete system includes a compiler-compiler, an ADT'81 compiler, an optimizer, a decompiler and an executer, and provides a complete class-based integrated platform for GIS design and development (Bouillé, 1994b).

3. INTRODUCING CONSTRAINTS

Any type may be associated to a condition, which looks like a string, which may be updated at run-time, and which will be executed any time the a.d.t. is invoked by any process. A typical example concerns an atribute named temperature; for creating a class, its attribute and specifying this one, the following statements have been used (figure 10); the result graphicaly appears on the figure 11.

```
CREATE "AREA" ;
.....
CREATE "AREA" . "temp" ;
.....
SPECIFY "AREA" . "temp" FLOAT UNIT (kelvin);
.....
"AREA" . "temp" ' COND := THIS."temp" > 0. ";
.....
```

Fig. 10 : Basical condition in ADT'81.

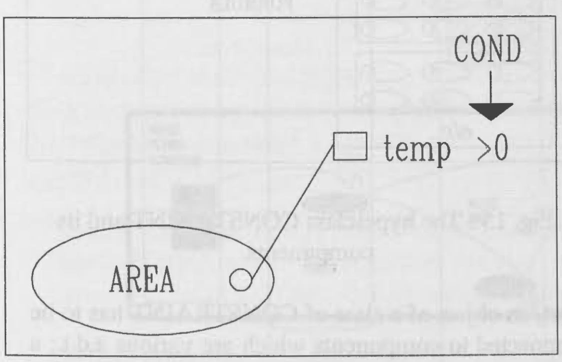


Fig. 11 : The consequence of the statements.

In this first simple example, an attribute only was concerned; in a second very simple example, we consider an area which number of neighbors cannot be greater than 6, for instance; for expressing that, we have created the link between classes which is "NEIGHBOR", it is a link on the class "AREA" itself; for obtaining at run time all the links going out of an object of the class "AREA", we need to use a functor entitled : OPSCRC (Object Positive Semi-Cocircuit restricted to the Relationship and final Class), and for obtaining the final objects, we need the functor entitled OFIN (Object FINal); for having the number of these neighbors, we just need the functor entitled CARD, which name naturaly does not require any explanation: the necessary statements are expressed on the figure 12 below.

```
CREATE "AREA" "neighbor" "AREA REFLEXIVE ;
.....
"AREA" "neighbor" "AREA" ' COND :=
THIS'OPSCRC("neighbor","AREA")'OFIN'CARD>+6 ;
.....
```

Fig. 12 : Basical condition on a link.

The result may be seen as expressed on the figure 13, expressing that the neighbor is carrying a condition which will be tested any time a new link between objects will be instanciated, corresponding to this link on the class "AREA".

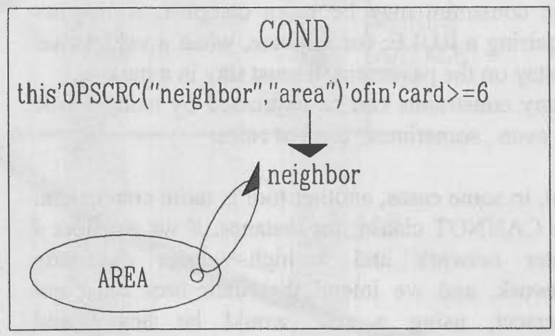


Fig.13 : The consequence of the statements.

This concept of condition, expressed by COND is not the only one allowing to perform automaticaly something in connection to a given a.d.t. ; another one, more complex is named ACTION.
An ACTION is a set of statements, associated to an a.d.t., accepting any programming statements, in any number; at run-time, when the a.d.t. is invoked, the action is automaticaly executed. This ACTION is a string which can be assigned, modified or deleted at run-time; for instance, it can be read from the user peripheral. Likewise for the condition. An action may for instance update the condition of any other a.d.t.. The User may easily remark that an a.d.t. simultaneously carrying a condition and an action exactly looks like a rule. In fact, it allows to make an economy of this one, and is more flexible. A condition, COND, may contain a more complex

boolean expression; it may contain too a fuzzy expression; every time, the condition only concerns the carrier, namely the a.d.t. carrying that condition; the real world, unhappily, presents conditions more complex, simultaneously linking several components, and in our approach, several a.d.t. That is why, we introduce new elements of vocabulary.

When a link between classes is created, carrying a relationship, for instance:

CREATE "ARC" "going out of" "VERTEX"
REVERSE "entering" ;

we express that an arc *MAY* go out of a given vertex; when creating a link such as "neighbor" :

CREATE "AREA" "neighbor" "AREA"
REFLEXIVE ;

we express that an area *MAY* have some neighbors.

If we want that a relationship MUST be verified, we express that in a different way; for instance, considering a vehicle A which has to follow another one, a vehicle B, in an urban network, we may use the next statement:

```
A [ "VEHICLE" ] MUST "follow"
            B ["VEHICLE"] ;
```

The constraint may be more complex, sometimes requiring a RULE; for instance, when a vehicle has to stay on the pavement, it must stay in a parking. Many constraints can be expressed by using a rule or even, sometimes, a set of rules.

But, in some cases, another tool is more convenient: the CANNOT clause; for instance, if we consider a water network and a high-voltage electricity network, and we intend that their arcs must not intersect, using a rule would be heavy and unelegant. To express that, simply concerning two arcs A and B, belonging respectively to the both networks, we only need the following statement:

```
A [ "MEDIUM VOLTAGE : ARC" ]
     CANNOT "intersect"
B [ "WATER : ARC" ] ;
```

In some cases, the relationship is a little more complex, and the CANNOT clause is not suficient, no more than a RULE. A classical example deals with the thermodynamics, where P (Pressure), V (Volume) and T (Temperature) are linked by a relationship. One only parameter cannot decide of the others; with two parameters, the third one is completely determined; instead of expressing three rules or emitting the formula and calling it every time, we only have to apply the concept of CONSTRAINT. For instance, if we have a DAMP, a TRANSFORMER STATION, some PILES supoporting a CABLE, and we want to protect a viewpoint in a PANORAMA, we have a complex (fuzzy) constraint which is represented on the figure 14 below (we do not give the statements).

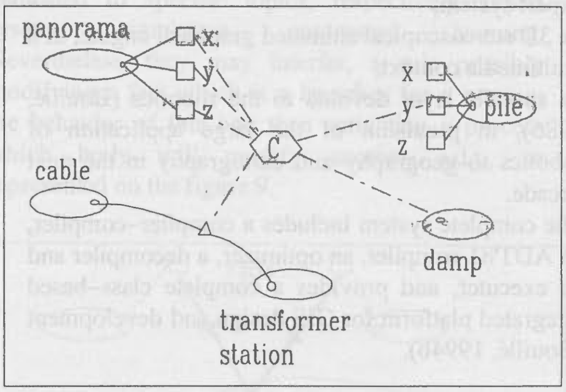


Fig. 14 : An example using the CONSTRAINT.

For dealing with constraints, we create a hyperclass named CONSTRAINT, which carries, among others, an attribute named FORMULA; all classes included in this hyperclass inherit of the FORMULA, which can be expressed in an algorithmical way, the structure being rather simple, shown on the figure 15 below. This body may contain things such as relationships belonging to the preceding cases: MAY, MUST, CANNOT, linking the constraint to other components.

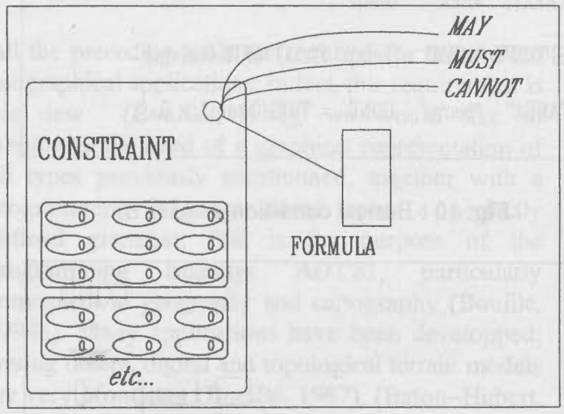


Fig. 15 : The hyperclass CONSTRAINT and its components.

But, an object of a class of CONSTRAINT has to be connected to components which are various a.d.t.; a link cannot join any type. According to a transformation based upon the theory of categories (Cousin, 1988), any a.d.t. may be associated with another different, allowing to build these necessary links. In toponymy automated positionning (Titeux, 1989), as well as in urban network structuring, the constraint is the mowt well-suited tool.

In the following example, shown on the figure 16, we want to express that any arc of the water network cannot intersect any arc of the medium–voltage network. Of course, it requires some computation in 3–D in order to determine the superimposition of arcs, and the distance between the two superimposed auxiliary points. The structure of figure 15 uses simultaneously the concepts of prototype, of structure, and of constraint.

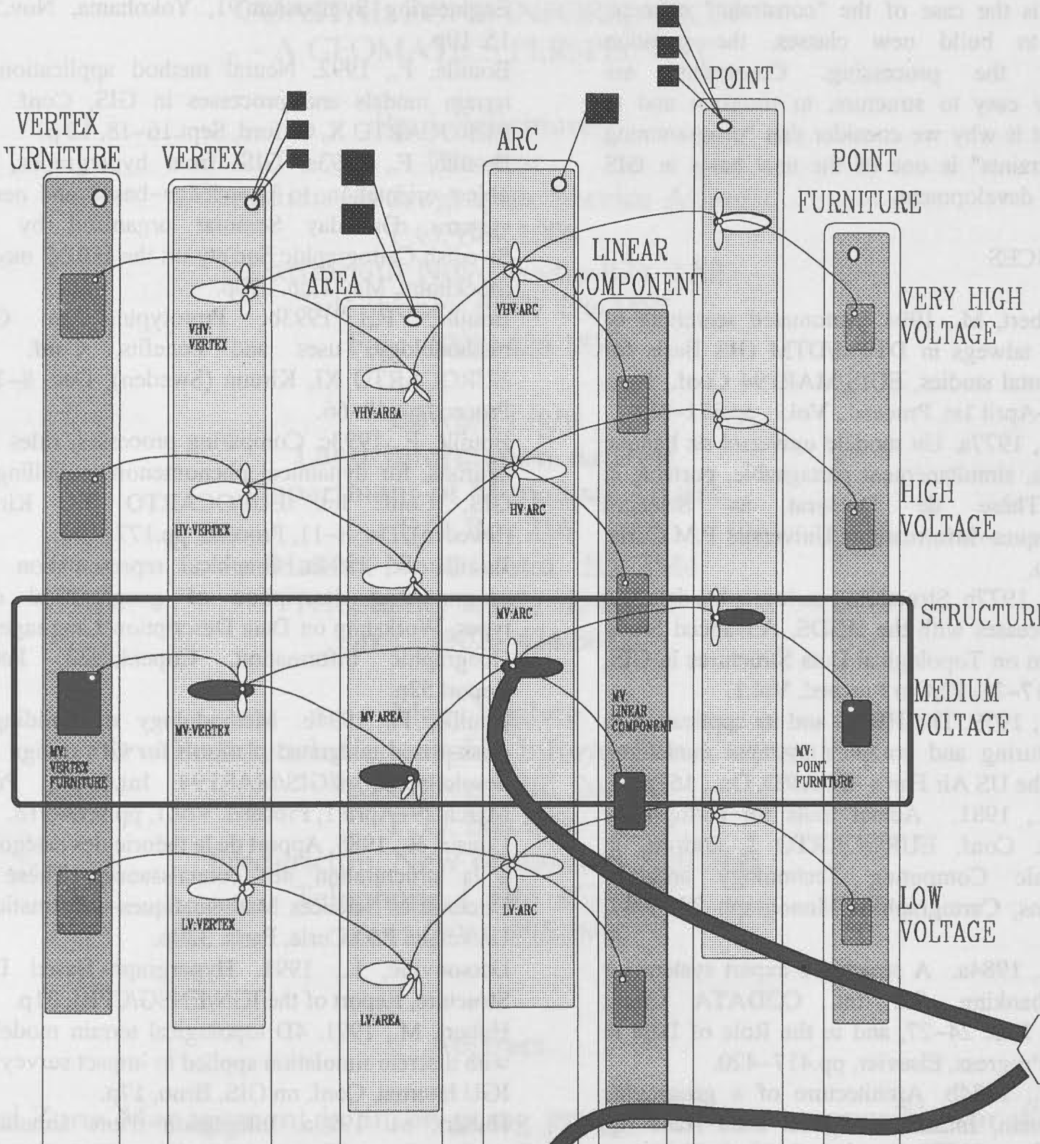


Fig. 16 : A complex constraint working on two networks.

4. CONCLUSION

The development of object-oriented GIS is increasing; automated processing of complex problems, which could not be foreseen before, can now be solved by applying new tools very easy to use. That is the case of the "constraint" concept, allowing to build new classes, the partition optimizing the processing. Constraints are particularly easy to structure, to program and to apply. That is why we consider that "programming with constraints" is one of the new basis in GIS design and development.

REFERENCES

Baton-Hubert, M., 1994. Automated searching of crests and talwegs in DTTM/DTM GIS Basis for environmental studies, EGIS/MARI'94 Conf., Paris March 29-April 1st, Proceed., Vol.1, pp.371-381.

Bouillé, F., 1977a. Un modèle universel de banque de données, simultanément partageable, portable et réparti, Thèse de Doctorat ès Sciences Mathématiques-Informatique,Université P.M.Curie, Paris, 550p.

Bouillé, F., 1977b. Structuring cartographic data and spatial processes with the HBDS, Advanced Study Symposium on Topological Data Structures in GIS, MIT, Oct.17-21, 22p. in Proceed. Vol.1.

Bouillé, F., 1979. The HBDS and its application to data structuring and complex systems modelling, Report to the US Air Force-EOARD, Oct., 167p.

Bouillé, F., 1981. Actual tools for cartography today, Int. Conf. EUROCARTO I, Oxford, in Cartographic Computing Technology and its Applications, Cartographica Monograph 28, 1982, pp.27-32.

Bouillé, F., 1984a. A structured expert system for geodata banking, 9th Int. CODATA Conf., Jerusalem, June 24-27, and in the Role of Data in Scientific Progress, Elsevier, pp.417-420.

Bouillé, F., 1984b. Architecture of a geographic expert system, Int.sem.on Spatial Data Handling, Zürich, August 20-24, Proceed., 24p.

Bouillé, F., 1986.Interfacing cartographic knowledge and robotics, Int. Conf. AUTOCARTO, London, Sept.14-19, Proceed.Vol.1, pp.563-571.

Bouillé, F., 1987. A survey on the HBDS methodology applied to cartography and land planning, Int. Conf. EUROCARTO VI, Brno , Proceed.pp.20-28.

Bouillé, F., 1988. Developping strategies in GIS, by problem-solving methods based on a structured expert system, Int. Conf. EUROCARTO VII, Enschede, Sept.19-22, Proceed. 12pp. and in Environmental Applications of Digital Mapping, ITC.Pub., N°8, pp.42-50.

Bouillé, F., 1991a. Decisive criteria for GIS comparative study and evaluation, IGU Int. Conf. on GIS, Brno (Tchecoslov.), 10p.

Bouillé, F., 1991b. Fuzzy neural processing by an object-oriented expert system - Application to Geographic Information Systems, Int. Fuzzy Engineering Symposium'91, Yokohama, Nov.13-15, 10p.

Bouillé, F., 1992. Neural method application to terrain models and processes in GIS, Conf. Int. EUROCARTO X, Oxford, Sept.16-18, 12 p.

Bouillé, F., 1993a. GIS: from hypergraphs, via object orientation, to knowledge-based and neural systems, One-day Seminar organized by the Swedish Cartographic Society on the HBDS model, Stockholm, May 12th. , 50p.

Bouillé, F., 1993b. Prototyping in GIS: methodology, uses and benefits, Conf. Int. EUROCARTO XI, Kiruna (Sweden), Dec. 8-11 , Proceed.pp.48-66.

Bouillé, F., 1993c. Comparing processes, rules and neurons, for dynamical phenomenon modelling in GIS, Conf. Int. EUROCARTO XI, Kiruna (Sweden), Dec.8-11, Proceed. pp.177-194.

Bouillé, F., 1994a. Graphical representation and programming description of geographical data types, Workshop on Data Description Languages in Geographic Information, Copenhagen, Feb.4, Report 52p.

Bouillé, F., 1994b. Methodology of building a class-based integrated platform for GIS design and development, EGIS/MARI'94 Int.Conf., Paris, March 29-April 1, Proceed. Vol.1, pp.909-918.

Cousin, R., 1988. Apport de la théorie des catégories à la structuration des connaissances, Thèse de Doctorat ès Sciences Mathématiques-Informatique, Université P.M.Curie, Paris. 320p.

Dassonville, L., 1991. Hypergraph-Based Data Structure, Report of the IGN/ENSG/CPRI, 31p.

Hubert, M., 1991. 4D topological terrain modeling with discrete simulation applied to impact surveying, IGU Internat. Conf. on GIS, Brno, 17p.

Hubert, M. 1993a. Intégration d'une simulation spatio-temporelle à un modèle topologique et numérique de terrain, Thèse de Doctorat Mathématiques-Informatique, Université P.M. Curie, Paris, 781p.

Hubert, M., 1993b. Expert system managing a large set of discrete processes for 3D cartography and mapping, Internat. EUROCARTO XI Conf., Kiruna, Dec.8-11, Proceed. pp.67-77.

Titeux, P., 1989. Automatisation de problèmes de positionnement sous contraintes: une méthode intégrant des techniques de système expert, de compilation et de bases de données. Application en cartographie, Thèse de Doctorat Mathématiques-Informatique, Université P.M.Curie, Paris, 414p.

APPROACHING THE MANAGEMENT OF COASTAL ZONE INFORMATION - A GEOMATICS PERSPECTIVE

Paul Bellemare
Director
Canadian Hydrographic Service Atlantic
P.O. Box 1006
Dartmouth Nova Scotia B2Y 4A2
Telephone: (902) 426-3497
Fax: (902) 426-1893

Edward Light
Land Information Analyst
Department of Municipal Affairs
P.O. Box 216
Halifax Nova Scotia, B3J 2M4
Telephone; (902) 424-3761
Fax : (902) 424-5872

Mary Ogilvie
New Brunswick Geographic Information Corporation
985 College Hill Road
P. O. Box 6000
Fredericton, New Brunswick E3B 5H1
Telephone (506) 453-3908
Fax: (506) 453-3898

Abstract

Coastal Zone Management is an on-going process. Those responsible for it's effective management, require accurate and accessible information. Through the development and implementation of land and water information standards the geomatics industry offers coastal zone manager the means by which reliability and integrity can be maintained for such information management.

Approche géomatique de la gestion de l'information relative aux zones côtières

Résumé

La gestion des zones côtières est un processus continu qui présuppose l'accès à des systèmes d'information pouvant fournir des données exactes. En élaborant et en mettant en application des normes pour l'information sur les terres et les eaux, l'industrie de la géomatique offre aux responsables de la gestion des zones côtières un moyen d'assurer la fiabilité et l'intégrité de la gestion de l'information.

INTRODUCTION

In the last five years there has been increasing interest in information about the coastal zone. The accident with the Exxon Valdez; the Green Plan, published by the federal department of the Environment; the report of the Committee on Land Use and the Rural Environment, released by the New Brunswick Government in April, 1993, the development of aquaculture; and the growing conflicts in the use of the marine environment; all point to the increasing urgency felt by many levels of government to better equip themselves to manage the coastal zone.

Departments at both the federal and provincial level undertake data collection of land based data and water based data. This data is collected to support the mandates in those departments to manage resources. Private sector interests (oil industry, environmental consultants) also collect data, in order to make good business decisions, or to meet government requirements.

Coastal data is a combination of water based data and land based data. There is momentum growing at all levels to better coordinate this collection of information, and to better exploit the potential of that already collected. The management of this data is changing in response to financial, environmental, and regulatory needs that did not exist 10 years ago. This paper will review the main differences between water based and land based data, will discuss the importance of standards to the proper exploitation of the value of the data already collected, and will review several projects underway to improve efficiencies in the collection of data and its management, focusing on recent efforts in Atlantic Canada to cooperate better at the federal, provincial and private sector levels in order to streamline the management of coastal data.

The Canadian Hydrographic Service Atlantic is the collector of bathymetric information for the Atlantic Region. It is responsible for producing a range of products in order to insure safe navigation in Atlantic waters. The New Brunswick Geographic Information Corporation is a Crown Corporation charged with, among other things, providing basic geographic information for the province. The Land Information Management Service is a section of the Department of Municipal Affairs in Nova Scotia. It is charged with a range of objectives including the development and promotion of efficient collection, use and management of geo-referenced information.

DIFFERENCES BETWEEN LAND BASED AND WATER BASED DATA

There are some fundamental differences between land and water based data. It is important to understand these differences in order to appreciate the challenges in trying to merge them in order to create a dataset suitable for managing coastal information.

Firstly, the physical reference for vertical information in the two datasets is fundamentally different. Land based data is referenced above a vertical reference point (normally determined by tide gauge data from several points). Bathymetric information, while tied to the same points is referenced to a local datum and the absolute value changes with the effect of water level fluctuations . Horizontal reference surfaces are essentially the same, although the precision with which horizontal position is determined is often higher for land based data than for water based data.

Time is a very important element of much of the data collected at sea - such things as water temperatures, salinity, fish landings, chemical concentrations etc. vary with time of year, time of day, and vertically within a column of water. Although time may be a factor in some land based data, it generally does not play as prominent a role.

The techniques available for data collection are different for collecting land based data as compared with water based data. It is easy to "see" the objects you want to map on the land - this makes aerial photography, GPS positioning, remote sensing from satellites, and traditional ground surveys appropriate methods for collecting this data. It has not been as easy to "see" what is under the water's surface. Recent advances in swath and sweep techniques have made it easier to "see" the sea bottom, but most bathymetric data available today is still of the sounding profiles variety, and most water based thematic data is still being collected as points, leaving large areas of uncertainty filled with interpreted estimates.

The water based data community has concentrated recent research dollars in two main areas. One is to devise a database engine, with facility for handling spatial temporal information and even multi dimensional information that could be used in a distributed environment. The other is to coordinate the collection and management of water-based data from all sources. Management of information is clearly a priority with the water based custodians. In general, recent efforts managing the land based data has been more in the applications area - devising new ways to use the data already collected to produce a new service for a customer.

Developing a coastal dataset - one which merges data taken from these two worlds - thus poses some challenges. Standards are important. At the very least standards dealing with reference systems, data models, access to information etc. must be pursued. Cooperation among the different players is equally important. The effort required should not be underestimated - much of the entire water based community and much of the land based community must be involved. There has been significant progress in achieving cooperation between the feds (CHS) and the provinces (NS and

NB) in Atlantic Canada. This is essential if good coastal information is to be provided to the managers in Atlantic Canada. The promise of some of the emerging technologies for near-shore bathymetric collection and for data management may make the job easier than it would have been a decade ago.

STANDARDS

Land information standards have been the subject of considerable activity over the past 4 or 5 years. New Brunswick has a committee, Nova Scotia has a committee, and nationally the Canadian General Standards Board (CGSB) committee on Geomatics have been active in discussing appropriate standards for the land based data. Both provincial committees are incorporating water information standards in their deliberations. Success in reaching consensus in these different committees has been mixed, but dialogue has been opened, and agreement is some areas has been reached. The interests of the International community, normally expressed through NATO, have been aired in the discussions of land based standards, but they are not the driving force.

On the other hand, most of the activity in water based standards is at the international level. Canada has been a major driving force at the International Hydrographic Organization (IHO) over the last 10 years in setting standards for water based information, and has driven many of the decisions and recommendations with respect to datum, accuracy standards, and exchange formats.

Coastal Information Standards must make the most of the standards already in place for water based data and for land based data, and must try to reconcile differences where they exist. In Atlantic Canada, the vehicle for discussion of coastal information standards has been a working group of the Atlantic Coastal Zone Information Steering Committee (ACZISC). So far agreement has been reached on standards for vertical datum and for horizontal datum, and work has begun on creation of a common digital coastline.

COOPERATIVE PROJECTS

There are many multiple-partner-projects in the Atlantic Region dealing with the coastal zone. Examples are

- Coastal Information Technology Architecture Program
- East Coast of North America Strategic Assessment Program
- Gulf of Maine
- ACAP
- Mahone Bay
- Electronic Chart
- Bouctouche Bay

- Shelburne
- CARP
- ICOIN
- Passamaquoddy ICOIN
- MacLaren Plansearch/AXES project
- Atlantic Coastal Zone Information Steering Committee
- Atlantic Coastal Zone Database Directory

Although there is much cooperative activity, there are problems in having a plethora of coastal projects. There is an increased emphasis being placed on community-based projects (i.e. ACAP sites, Bouctouche Bay, Mahone Bay). Resources are more likely to be allocated to these projects rather than to global projects. Many of the more expensive projects are funded on the basis of a little money from a lot of people. This has appeal in that for a relatively small investment an agency gets the benefit of a very comprehensive project. However, with the number of large projects needing funding, smaller jurisdictions are looking at multiples of a small amount of money, and this adds up. Outside the coastal community, and perhaps to some extent within it, there is confusion about the goals and deliverables of the different projects, and how they differ from one another. This confusion leads to difficulties in getting funding for the large projects.

FEDERAL-PROVINCIAL COOPERATION IN ATLANTIC CANADA

The Atlantic Coastal Zone Information Steering Committee is a forum for federal and provincial governments in the region to discuss their activities, and to work cooperatively on projects that are of interest to all. The Atlantic Coastal Zone Database Directory is one of these projects. It is a compendium of meta data about databases dealing with the coastal zone, collected from provincial and federal agencies in the region. It is available on hard copy or on diskette form. Coastal Zone 94 is another. This working group has been preparing for a conference on the coastal zone to be held in Halifax in the fall of 1994. Working groups of ACZISC also deal with standards, and with several of the cooperative projects listed earlier in this paper.

On a bilateral basis, there has been provincial-federal cooperation on exchange of digital data, and on creating a common coastline for users both federal and provincial.

CONCLUSIONS

Proper management of coastal zone information is critical. There is public pressure to manage the coastal zone much better than it has been in the past, and our political masters will insist that we find mechanisms to make this happen. Although this paper has highlighted some of the problems in working in a cooperative way, these problems must be resolved. Certainly the experience in Atlantic Canada shows that cooperative problem solving is possible.

LE PLAN GÉOMATIQUE GOUVERNEMENTAL AU QUÉBEC, UNE STRATÉGIE POUR UNE PLUS GRANDE EFFICACITÉ DE L'ADMINISTRATION PUBLIQUE

Francine C. Boivin

Secrétaire générale associée
Secrétariat à l'aménagement, au développement régional et à l'environnement
Ministère du Conseil exécutif
Gouvernement du Québec

A Governmental Geomatics Plan for Quebec:

A Strategy for a More Efficient Public Administration

MOTS-CLÉS : Plan géomatique, décloisonnement, révision des façons de faire, échange de données.

ABSTRACT

In 1988, the Government of Québec has created a "three-level structure" in order to meet the increasing needs for coordination in Geomatics : A Committee of Deputy Ministers, a Steering Committee on Geomatics and a Users' Forum. In 1992, the Government has adopted a comprehensive plan aimed at developing Geomatics in the different governmental departments concerned with land management. Two priorities can be identified from the actions taken since then : (1) The need for the exchange of georeferenced data between different organizations; and (2) the management of tasks re-engineering in the public administration. In that perspective, accurate training will have to be given to every manager to ensure that geomatics is properly implemented in a context of limited financial resources.

Voyons d'abord un bilan des résultats obtenus jusqu'ici par l'effort d'équipe fourni depuis 1989 par la structure de coordination mise en place pour optimiser les investissements du gouvernement québécois en géomatique. L'exposé traitera ensuite du décloisonnement des données pour appuyer les échanges entre les ministères ainsi que la révision des façons de faire. En terminant, les nouvelles pratiques de gestion intégrée du territoire suscitées par la géomatique seront abordées. Ceci devrait permettre d'observer comment, par des projets concrets, le Québec progresse sur la voie de l'efficacité organisationnelle.

I- UN PREMIER BILAN

En novembre 1988, le Conseil des ministres confiait au Secrétariat à l'aménagement, au développement régional et à l'environnement (SADRE) le mandat de coordonner les activités touchant la géomatique. L'importance de pouvoir rendre accessible à tous l'abondante quantité d'informations caractérisant le territoire appelait une intégration et une normalisation de celles-ci.

Aucun ministère ne pouvant seul couvrir tous les aspects de l'implantation de la géomatique, une structure de coordination à trois niveaux s'est mise en place dès 1989. D'abord le Comité des sous-ministres de la géomatique définit les enjeux majeurs et les orientations à suivre. Pour sa part le Comité directeur voit à l'élaboration et à la mise en oeuvre du Plan géomatique gouvernemental. Enfin, le Forum des usagers apporte au Comité directeur l'éclairage des spécialistes qui ont à transformer leur environnement organisationnel en fonction du processus de géomatisation (voir annexe I).

Les premiers travaux exploratoires réalisés entre 1989 et 1991 ont permis de préciser les balises qui permettent de rationaliser les investissements en géomatique, notamment par une stratégie d'échange de l'information entre producteurs et utilisateurs. Dès juillet 1990, une enveloppe budgétaire a été accordée au Comité directeur de la géomatique pour faciliter des actions structurantes dans ce domaine.

En juillet 1992, le Conseil des ministres confiait au SADRE le mandat de mettre en application le **Cadre de référence pour la mise en oeuvre du Plan géomatique gouvernemental**. Pour ce faire, un plan d'action triennal a été élaboré couvrant les exercices 1992-1993, à 1944-1995. Les orientations ayant été bien définies, il s'agissait de passer à l'action pour traduire, dans les opérations, la volonté gouvernementale d'obtenir une plus grande productivité dans l'accomplissement des missions essentielles de l'État sur le territoire.

La stratégie du gouvernement définie par ce plan d'action reposait sur les principes de base suivants:

•concevoir chaque initiative ou projet d'investissement en géomatique dans une perspective gouvernementale afin de faciliter les échanges de données à référence spatiale entre les ministères, de réaliser plus efficacement les missions gouvernementales de nature territoriale et d'optimiser les modes de fonctionnement des unités administratives concernées;

•reconnaître que le complètement, par le ministère des Ressources naturelles de la base des données géographiques et foncières (BDGF), constitue une activité prioritaire;

•respecter les orientations gouvernementales àl'égard des technologies de l'information et retenir à cet effet que seuls seront approuvés les projets sectoriels en géomatique qui démontrent leur rentabilité ou leur caractère de nécessité pour réaliser une mission gouvernementale décrétée par législation;

Le gouvernement décidait également de consolider les structures de coordination permettant de mettre en commun les compétences et de favoriser des approches interministérielles en matière de gestion, d'exploitation et d'aménagement du territoire.

Une enveloppe budgétaire gérée centralement par le SADRE au nom du Comité directeur de la géomatique a été maintenue jusqu'au 31 mars 1995.

Le Comité directeur de la géomatique, là où l'effort d'équipe se construit.

Le Plan géomatique gouvernemental ne pourrait atteindre ses objectifs sans la collaboration de tous. Le Comité directeur regroupe les représentants des ministères les plus actifs et directement concernés par la géomatique. À travers cette structure de coordination qui se veut souple et basée sur l'établissement de solides consensus, le décloisonnement des données entre producteurs devient possible. Des liensfonctionnels entre plusieurs unités administratives s'établissent permettant alors un véritable effort d'équipe. Ceci permettra de créer l'environnement administratif nécessaire pour passer à l'étape de la révision des façons de faire au cours des prochains mois.

Rappelons que le Comité directeur de la géomatique est composé des représentants des ministères et organismes suivants :

- Le Secrétariat à l'aménagement , au développement régional et à l'environnement du ministère du Conseil exécutif;
- Le Secrétariat du Conseil du Trésor;
- Le ministère des Ressources naturelles;
- Le ministère des Transports;
- Le ministère de l'Environnement et de la Faune;
- Le ministère des Affaires municipales;
- Le ministère de l'Agriculture, des pêcheries et de l'alimentation.

II- UNE PRIORITÉ: LE DÉCLOISONNEMENT DES DONNÉES POUR APPUYER LES ÉCHANGES

Les ministères font face à un défi de gestion. Il est de moins en moins possible de réduire les dépenses publiques par de simples compressions budgétaires. Il faut maintenant repenser la façon de traiter avec les citoyens et leur fournir des services au moindre coût possible. Il faut s'assurer que les structures administratives offrent la souplesse nécessaire pour accroître la productivité. Dans le contexte actuel, le Comité directeur de la géomatique se doit d'apporter une contribution significative. Il poursuit une stratégie de moyen terme pour mieux systématiser le traitement et l'échange de l'information territoriale, faciliter la prise de décision et rendre possible la concertation interministérielle dans l'analyse de situations complexes.

Concrètement, cela a donné lieu à plusieurs travaux visant à lever les derniers obstacles à la libre-circulation des données à référence spatiale. La normalisation du format d'échange des données est certes un domaine où les spécialistes du Québec ont eu de nombreux contacts avec les chefs de file en ce domaine au Canada.
D'importants travaux sont en cours pour identifier le scénario le plus avantageux afin de normaliser le format d'échange de l'information à référence spatiale. Ces travaux sont faits en tenant compte des consensus qui sont en train de s'établir sur cette question autour du OPEN GIS Foundation. L'Université Laval, le Secrétariat du Conseil du Trésor et le ministère québécois des Ressources naturelles participent avec l'OGIS à la progression d'un tel consensus nord-américain.

Le Comité directeur de la géomatique a aussi favorisé le décloisonnement des banques de données de façon à assurer une plus grande homogénéité et complémentarité entre elles. Ce faisant, c'est toute la chaîne de montage de l'information territoriale qu'il s'agit de constituer à partir des multiples productions de données territoriales dispersées à travers les ministères et organismes. Le développement du géorépertoire documentaire viendra compléter cet effort. Ce géorépertoire se traduira par un logiciel (METASIRS) qui viendra s'arrimer à tout système d'information pour caractériser la qualité de l'information qui le supporte. L'utilisateur d'un système d'information àréférence spatiale (SIRS) saura alors exactement comment l'information a été saisie, si elle est structurée pour fins d'analyse spatiale, quel est son cycle de mise à jour, son niveau de précision, etc.

- Établir des liens entre les ministères

Au cours de 1994, des tests seront réalisés pour évaluer la capacité de réaliser des échanges de données à référence spatiale en mode numérique à partir d'un réseau de communications informatiques. Ainsi, il sera possible de tirer avantage de l'infrastructure électronique qui supportera la circulation des données entre les ministères. Cette infrastructure sera mise en place progressivement au cours des prochaines années de façon à permettre une diffusion plus élargie de l'information territoriale.

III- LA RÉVISION DES FAÇONS DE FAIRE

Le décloisonnement des données ne règle pas tous les problèmes. Il est nécessaire de passer à un second niveau de décloisonnement pour que la géomatique puisse utiliser à son mieux ou rentabiliser sa capacité d'intégration de l'information. Il s'agit du décloisonnement des structures administratives. C'est à ce niveau que le rôle du gestionnaire devient essentiel et déterminant. Pour être efficace, le gestionnaire doit s'impliquer pour évaluer la meilleure combinaison possible de ressources pour concourir à un résultat précis. Ce n'est qu'en fonction d'un plan stratégiquedifinissant les objectifs de chacun des ministères que la géomatique pourra être mise à contribution dans la mise en pratique des nouvelles façons de faire.

Comment appuyer les gestionnaires dans leur travail, eux qui sont les premiers confrontés aux exigences de l'efficacité ? La compétitivité d'une part et le poids de la dette publique d'autre part ne laissent aucune alternative: les processus administratifs doivent être réinventés, les structures de gestion repensées, les systèmes d'information réorganisés. Un premier outil logiciel a été développé à l'Université Laval à ces fins. Il porte l'acronyme PHENIX.

Par des moyens comme celui-ci le Comité souhaite que la géomatique soit introduite comme un volet important d'une stratégie nouvelle de gestion et non comme un simple apport technologique.

La révision des façons de faire est préconisée par plusieurs comme étant la seule manière d'opérer des changements crédibles et significatifs pour alléger et simplifier les processus de gestion. Trois axes d'application sont ici privilégiés. Ils fournissent une plateforme d'apprentissage aux gestionnaires qui pourront, par l'expérimentation de nouveaux procédés administratifs, procéder à cette révision des façons de faire.

- L'application du concept de gestion intégrée des ressources

L'objet de ces travaux consiste à éprouver un outil de simulation pour mesurer l'impact des interventions gouvernementales en milieu forestier. Déjà, un premier modèle de simulation a fait l'objet d'une validation dans le parc de la Mastigouche à l'est de Montréal. Ce modèle optimal d'intervention met en concurrence l'eau, le paysage, la faune et les ressources forestières. Il permet de mesurer, pour chaque scénario envisagé, les conséquences économiques et environnementales qui en découlent (voir schéma).

- La mise en oeuvre d'une approche globale de gestion du territoire

Il s'agit ici d'un processus qui permettra de faire la synthèse de toutes les interventions gouvernementales sur les terres publiques. Cette même approche sera également supportée pour assurer le suivi des interventions gouvernementales par rapport aux schémas d'aménagement des MRC et en relation avec l'application des droits territoriaux des peuples autochtones. Divers travaux seront progressivement mis en route qui permettront d'intégrer toutes les interventions, réglementations de zonage, limites administratives, droits d'occupation, usages divers qui se retrouvent dans une même entité territoriale. Cette approche interministérielle devrait contribuer à concilier les usages conflictuels entre les ministères sur un même territoire.

- L'application du concept de gestion intégrée par bassin

Au cours de 1994, le ministère de l'Environnement et de la Faune, de même que celui de l'Agriculture, des Pêcheries et de l'Alimentation, mettront en commun leurs efforts afin d'ajuster leurs programmes respectifs de développement durable à l'échelle du bassin hydrographique. Le ministère des Transports et Hydro-Québec auront à contribuer pour mettre en place une vision intégrée facilitant la conciliation des usages sur le territoire à l'échelle d'un bassin hydrographique. Ce concept n'est pas nouveau mais sa mise en application nécessite une coordination de plusieurs équipes de spécialistes et nécessite la convergence de plusieurs approches scientifiques.

Par rapport à cette démarche, le Comité directeur a fait appel à l'Université Laval pour mettre à l'épreuve le logiciel PHENIX. Ce logiciel est un outil qui permet de faire le diagnostic de la gestion des banques de données. Il permet de les cataloguer, d'établir la fréquence de leur mise à jour et d'analyser les redondances et les duplications dans la constitution des fichiers. Un inventaire de toutes les données hydriques et atmosphériques à référence spatiale a donc été produit. À partir de ce résultat il sera possible de planifier la réingénierie des processus de collecte et de traitement de cette information. Au cours des prochains mois, les ministères concernés par cette démarche procèderont à la normalisation de la définition numérique des périmètres des bassins hydrographiques. Ce n'est qu'après avoir franchi ces étapes qu'il sera possible d'appliquer efficacement des logiciels d'analyse de diffusion des polluants sur ces bassins hydrographiques.

IV- LA GÉOMATIQUE, UN PROCESSUS QUI PERMET L'EFFICACITÉ ORGANISATIONNELLE

Nous avons tous comme gestionnaires à faire preuve d'efficacité. La géomatique permet, nous l'avons constaté avec les trois axes d'application identifiés précédemment, de créer des liens fonctionnels entre plusieurs unités administratives. Un décloisonnement s'opère et chaque unité de production contribue à une partie de l'information pour réaliser une opération commune de nature interministérielle qui permet un effort d'équipe et de mise en commun des ressources informationnelles.

La géomatique permettra aussi l'intégration de données de sources diverses. Il sera possible de dresser une synthèse d'une situation complexe. Par une modélisation appropriée, des simulations seront possibles pour mesurer l'impact des interventions du gouvernement, que ce soit pour planifier l'exploitation des forêts, contrôler la pollution des cours d'eau, optimiser les investissements sur le réseau routier ou modifier les pratiques des agriculteurs pour accroître les rendements.

La géomatique est un processus de gestion qui s'allie à plusieurs technologies dont celle du GPS (Global Positioning System) et des télécommunications. Le décloisonnement des tâches est facilité par la rapidité avec laquelle l'information de base, captée par un carnet de note électronique sur le terrain, devient maintenant accessible auprès d'une foule d'utilisateurs. Les spécialistes et les gestionnaires sont donc de plus en plus reliés entre eux par une chaîne de montage électronique qui transcende les structures administratives et hiérarchiques. On assiste progressivement à un réseautage qui s'accentuera par le développement des liens électroniques entre les banques de données découlant de la stratégie de mise en réseau des réseaux du programme national de l'autoroute de l'information (CANARIE). Cette révolution dans la manière de travailler est à la base de ce que nous appelons la géomatisation des organisations.

Tout cela vient très vite. La technologie permet une foule de choses, elle supporte la transformation des organisations, des façons de faire, mais elle exige aussi beaucoup des personnes appelées à utiliser la géomatique dans un tel contexte. Il apparaît donc urgent et essentiel d'investir dans la formation des ressources humaines. Toutefois, le contexte budgétaire risque de freiner l'application de la géomatique. Il y aura donc des choix à faire pour privilégier certaines dépenses plus que d'autres.

Le Plan géomatique gouvernemental prévoit des mesures d'accompagnement pour assister les gestionnaires dans la géomatisation des organisations. L'expertise développée au Québec à travers les projets de recherche, les institutions d'enseignement supérieur et les contrats réalisés par les entreprises de services-conseils est actuellement l'objet d'une structuration. Les besoins en formation sont importants au chapitre de la géomatisation des organisations. C'est pour faire face à cette réalité et aux ressources budgétaires limitées qu'universités, entreprises et centres de recherche unifient présentement leurs efforts pour mieux relever ce défi.

La géomatisation des organisations est en fait la transformation des méthodes de travail et d'intervention sur le territoire grâce au support de la géomatique. Cette transformation qui s'opère au Québec contribuera, grâce au Plan géomatique gouvernemental, au maintien d'une Fonction publique efficace et soucieuse de sa compétitivité et d'espérer atteindre cette efficacité organisationnelle tant souhaitée.

CONCLUSION

Les éléments essentiels de l'évolution du mandat gouvernemental en géomatique au gouvernement du Québec ont été ici présentés. Le succès de cette vaste opération de coordination au plus haut niveau àlaquelle sont associés tous les usagers gouvernementaux de la géomatique tient en trois mots : coordination, échange et décloisonnement des structures. Nous pourrions y ajouter la participation de tous suscité par l'esprit de collégialité avec lequel les décisions sont prises au Comité directeur de la géomatique. La révison des façons de faire nécessite une remise en question en profondeur de plusieurs habitudes acquises au fil des ans par les gestionnaires. Le défi qui reste à relever est donc de fournir une formation continue à ces gestionnaires pour les rendre aptes à mieux contrôler ce nouvel environnement de gestion. Ainsi avec tous ces atouts en main il sera possible de procéder avec succès à la géomatisation de nos organisations.

Integration of GPS, Photogrammetry, and GIS

Capt. Lewis A. Lapine, NOAA
Chief, National Geodetic Survey
Coast and Geodetic Survey
National Ocean Service
National Oceanic and Atmospheric Administration
1315 East-West Highway
Silver Spring, MD 10910

ABSTRACT

The Photogrammetric Branch of NOAA now produces accurate base maps by combining the technological advances of aerial photography, Global Positioning System (GPS), and Geographic Information System (GIS). The combined technologies have enabled the production of a Florida Keys National Marine Sanctuary GIS. This composed of approximately 3500 square nautical miles of interconnected islands and benthic communities. Much of this area is being over-stressed by commercial development and recreational activities. A description of this project will be presented in its phases of operations: the initial field work, GPS assisted aerial photography and aerotriangulation, digital compilation, and benthic classification. The GIS as a final product will also be presented.

A futuristic look at the integration of other imagery sensors, such as Multispectral Scanners (MSS) in the performance of mandated programs will be discussed. Anticipated applications of MSS and photographic composite data sets include efficient inventory of wetlands, quick response for natural disaster damage assessment, pollution tracking, submerged aquatic vegetation mapping, and marine mammal studies. A marine resources GIS may play a central role in the production of maps and digital data available for traditional nautical charting, the developing industry of electronic charts and marine resource management.

KEY WORDS: Global positioning system, Photogrammetry, Geographic information system, Multispectral scanner, Marine resource management, Benthic classification, Nautical charting, Electronic charts.

INTÉGRATION DU GPS, DE LA PHOTOGRAMMÉTRIE ET DES SIG

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif

Civic Addressing for the County of Lennox and Addington

Rowland Tinline, Susan Holt, Carolyn Fielding

Queen's/IBM GIS Laboratory
Department of Geography
Queen's University
Kingston, ON K7L 3N6
Phone: (613) 545-6039
Fax : (613) 545-6122

ABSTRACT

This paper reports on how civic addressing was accomplished for approximately 10,000 properties in the County of Lennox and Addington in Ontario. The major features of this project included: (i) real time mapping of the road network using GPS and Field Notes (TM) software; (ii) real time assignment of addresses using "fifth wheel" technology on loan from Bell Canada; (iii) address verification using a combination of the tax rolls, Bell Canada's service address data base and Canada Post's route lists; and (iv) a large component of local volunteer help for the preliminary fieldwork. The paper describes each of those features and assesses their costs, problems and role in future civic addressing projects.

Établissement des listes de numéros civiques pour le comté de Lennox et Addington

RÉSUMÉ

Le présent document explique comment ont été établies les listes des numéros civiques d'environ 10 000 propriétés, dans le comté de Lennox et Addington, en Ontario. Les points saillants de ce projet comprennent notamment : (i) la cartographie en temps réel du réseau routier au moyen du système de positionnement global et du logiciel Field Notes (md); (ii) l'attribution d'adresses en temps réel au moyen de la technologie «surpuissante» prêtée par Bell Canada; (iii) la vérification des adresses par recoupement des données des rôles d'imposition, de la base de données des adresses de service de Bell Canada et des listes d'itinéraire de la Société canadienne des postes; (iv) la quantité appréciable d'information fournie par des volontaires des régions faisant l'objet de l'étude, en vue des travaux préliminaires sur le terrain. Le document décrit enfin chacune de ces caractéristiques et évalue leur coût, les difficultés qu'elles entraînent et leur importance en ce qui a trait à d'éventuels projets d'établissement de listes de numéros civiques.

INTÉGRATION DU GPS, DE LA PHOTOGRAMMÉTRIE ET DES SIG

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif

Precise Georeferencing and Classification of Airborne Pushbroom Multispectral Imagery

D. Cosandier, M.A. Chapman and P. Gong
Department of Geomatics Engineering
The University of Calgary
2500 University Drive N.W.
Calgary, Alberta
Canada, T2N 1N4
Tel.: (403) 220-5834; Fax: (403) 284-1890

ABSTRACT

Multispectral image data are finding increased application in areas such as environmental impact assessment, forest inventory management and change detection. A basic requirement for these and other applications is the need for precise georeferencing and classification. These requirements must be met in order for the data to play an effective role when integrated with other data types in geographic information systems (GIS) and decision support systems. This paper outlines a mathematical model which has been developed and implemented in software for the georeferencing of airborne pushbroom imagery. The current formulation incorporates auxiliary navigation data, particularly GPS and inertial navigation system (INS) data, for interpolating position and attitude information related to the imaging sensor exposure stations. Using the precisely georeferenced images, the classification of scene details is carried out using hybrid algorithms which have been tailored for specific applications. Results for test data sets are presented which illustrate the effectiveness of the georeferencing algorithm as well as the reliability of the classification methodology. This research was conducted under the sponsorship of a Natural Science and Engineering Research Council Strategic Grant.

Proposed for ISPRS Symposium, Commission II
Session: Real Time Mapping Technologies
Ottawa, June 1994

Géocodage précis et classification des imageries multispectrales aériennes obtenues de scanneurs à barette de détecteurs

RÉSUMÉ

On trouve de plus en plus d'applications aux données d'imagerie multispectrale notamment dans les domaines de l'évaluation des incidences environnementales, de la gestion de l'inventaire forestier et de la détection des transformations. Dans chacune de ces applications, on doit procéder au géocodage précis et à la classification des données recueillies afin de s'assurer qu'elles remplissent efficacement leur rôle lorsqu'elle sont intégrées aux autres données des systèmes d'information géographique (SIG) et des systèmes d'aide à la décision. Le présent exposé traite d'un modèle mathématique intégré à un logiciel pour le géoréférencement de l'imagerie aérienne obtenue au moyen de scanners à barette de détecteurs. Dans sa forme actuelle, ce modèle incorpore des données auxiliaires de navigation, notamment des données provenant des systèmes de positionnement global (GPS) et de navigation à inertie, permettant d'interpoler de l'information sur la position et l'attitude des points de prise de vues des capteurs-imageurs. Sur la base des images précisément géoréférencées, on effectue la classification des détails observés à l'aide d'algorithmes hybrides conçus en fonction d'applications particulières. On présente les résultats obtenus avec des ensembles de données échantillons. Ces résultats indiquent l'efficacité de l'algorithme de géoréférencement ainsi que la fiabilité de la méthode de classification. Cette recherche a été financée par une subvention stratégique du Conseil de recherches en sciences naturelles et en génie du Canada.

GUIDING SHIPS ON WATERWAYS

M. Sandler, U. Kabatek, R. Neul, E. D. Gilles

Institut für Systemdynamik und Regelungstechnik, Universität Stuttgart

Commission II, Working Group 1

KEY WORDS: ship guidance, radar image processing, sensor integration

ABSTRACT

A self-contained, integrated navigation system for the automatic guidance of ships on waterways is described. Besides the usual navigation sensors, imaging sensors like a radar or a laser scanner are used. The sensor information is combined with a-priori knowledge in the form of an electronic chart of the waterway and a dynamic ship model, in order to determine the position and heading of the navigating ship in global coordinates. Matching of image data and electronic chart is performed by a least-squares matching technique that is applied similarly to measurements from the different imaging sensors. The results of the matching processes and measurements from other sensors are integrated by a Kalman filter. A multiple-target tracking method is implemented for the determination of the trajectories of other vessels and thus for the evaluation of the actual traffic situation. Additional processing steps of the integrated navigation system are the planning of a trajectory and the guidance of the ship along this trajectory. The methods were successfully tested in practice. The structure of this integrated navigation system, developed for the case of inland shipping, may be transferred to other traffic systems.

1 INTRODUCTION

At the "Institute for System Dynamics and Control", University of Stuttgart, an integrated navigation system for inland and costal waterways is being developed. The goal of the project is a system capable of automatically guiding a ship on a waterway in normal traffic situations. Such a system will relieve the navigator from tiring routine work and support him in complex situations, particularly at night or in foggy weather. The heavy traffic on e. g. the river Rhine with shipment of dangerous goods such as chemicals and refinery products implies many risks for the environment. The system helps to protect the environment by increasing the level of safety on the waterway. It is not targeted as a replacement for the navigator.

This project is funded by the "Deutsche Forschungsgemeinschaft" within the SFB 228 'High precision navigation'. It is also promoted by the german ministry for traffic.

Within this paper we will focus on the real-time processing of the measurements of the different sensors, especially the processing of the image data. Image data are generated by a radar and by a laser scanner both sensing the surroundings of the ship. In section 2 the configuration and components of the navigation system will be explained first. The electronic chart is an important component in the image processing algorithms. Her structure and implementation is discussed in section 3. The image data are used to determine the position and heading of the own vessel in global coordinates by matching the images to the electronic chart of the waterway. The matching algorithms are discussed within section 4. As explained in section 5 the image data are also used to determine the actual traffic situation. Based on the matching results and measurements of other sensors available to the system an integrated estimate of the ship's state is computed by a Kalman filter described in section 6. In section 7 the generalized structure of the navigation system is deduced. Finally section 8 gives a brief outlook on the trajectory planning and control algorithms implemented within the navigation system.

2 COMPONENTS AND TASKS OF THE INTEGRATED NAVIGATION SYSTEM

The configuration of the integrated navigation system is shown schematically in figure 1. On the left side the sensors of the system are represented. Imaging and non-imaging sensors can be distinguished. The *radar* yields a map-like representation of the local environment. It is part of the standard equipment of most commercial inland and coastal ships. A new radar image is obtained about every 2.3 seconds and transferred into the ship-borne *computer*. Processing of the radar image is very computationally intensive and has to be accomplished in real-time. Therefore, all algorithms for this task have to be designed for max-

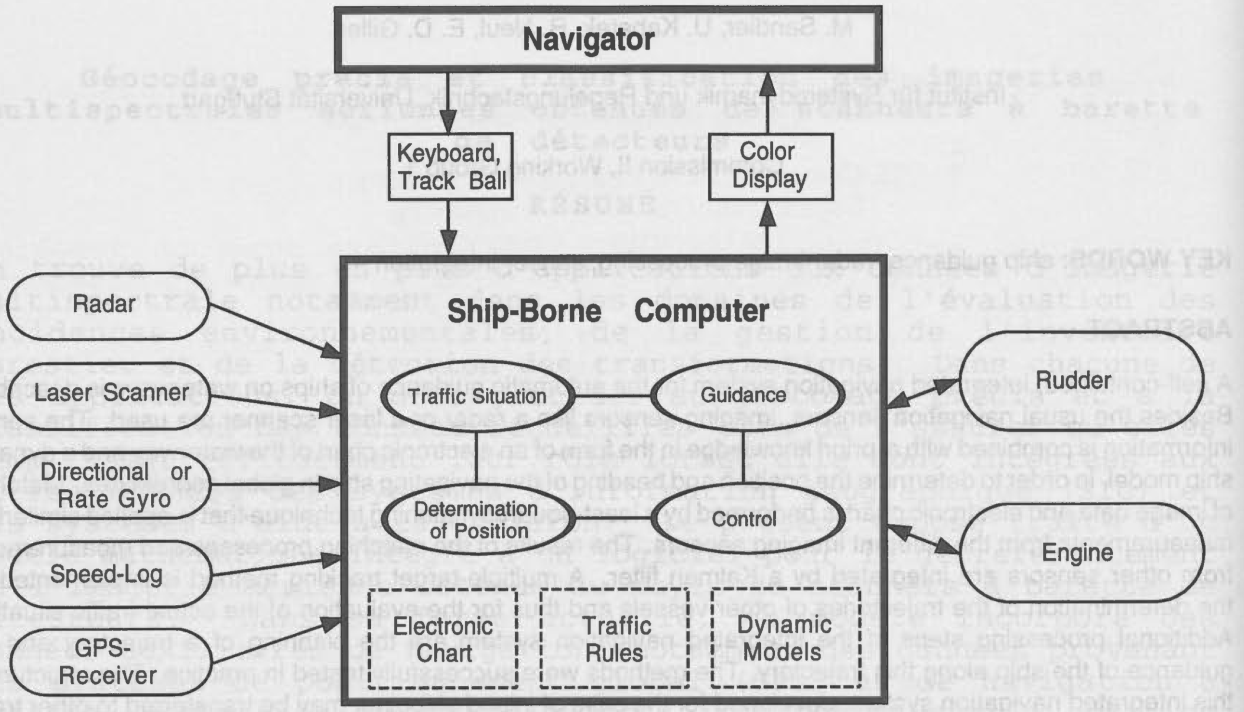


Figure 1: Configuration of the integrated navigation system

imum speed. A *laser scanner* will be used to enable automatic entering into locks and navigation in very narrow canals. The accuracy obtainable with a standard radar is not sufficient for this purpose. Among the non-imaging sensors, the *gyroscope* gives a measurement for the heading angle respectively for the turning rate of the vessel. The *speed-log* renders measurements of the speed over ground or relative to the water. The main output of the *GPS-receiver* consists of positional coordinates, based on an absolute coordinate system. It can also supply the speed over ground.

On the right side of figure 1, the actuators are shown. During an automatic cruise, the computer can control the *engine* throttle and the *rudder* of the vessel.

In the center of the figure, the navigation system is represented with all its tasks in the ship-borne computer. The computing hardware is a standard DEC-MicroVAX under the VAXELN real-time operating system. The first task, symbolized by an ellipse, is the *determination of the position* of the vessel. For this purpose, all the signals of the various sensors are combined in order to obtain the position and heading of the vessel. The matching of the radar image to the electronic chart, as explained in section 4, is a peculiar feature of this system.

The radar image is also utilized for the determination of the *traffic situation*. The trajectories of other vessels are tracked over sequences of radar images. Principally, every echo in the radar image located in the waterway can stem from a foreign vessel. A multiple-target tracking algorithm results in estimates for the position, speed and heading of these vessels.

The results of the determination of the position and the traffic situation are fed into the third task of the navigation system, the *guidance* of the vessel. The trajectory planned for the vessel is input to the *control* task. This task acts on the rudder and engine throttle of the vessel.

The integrated navigation system is supervised by the human *navigator* of the vessel. The results of the different tasks are presented on a *color display*. The navigator can interact with the system by means of a *track ball* and a *keyboard*, supplementing or overriding the results of the guidance task.

A major resource for all the tasks of the integrated navigation system is the a-priori knowledge deposited in knowledge bases within the computer. These knowledge bases are displayed as dashed boxes in figure 1. They consist of a database for the *electronic chart*, *traffic rules* and *dynamic models* for the own craft and for other vessels. The knowledge bases contain most of the information the navigator himself has to possess for navigation purposes. For example, the electronic chart supplies the route knowledge to the navigation system.

The navigation system has been tested successfully in numerous trials. The test have been carried out at the institute's test vessel "Falke" (16 m, 15 t), the measuring ship "Neckar" (30 m , 200 t) and the commercial vessel "Neuenstein" (105 m, 1900 t). Recently a first test on the push tow "L16" (total of 185 m, 10000 t) has been undertaken.

3 THE ELECTRONIC CHART

3.1 Requirements of the electronic chart

The structure of the electronic chart conforms to the distinct requirements within the integrated navigation system. The format of the database for the storage of the electronic chart is determined mainly by the demands of real-time processing. Furthermore, the comparison of the chart with the radar image requires a specific type of spatial access to the data of the electronic chart. This access sets out from a point given in absolute coordinates. The database returns the nearest objects in the chart and their Euclidian distance to these input coordinates. The distance is calculated as the minimum of the distances to all points on the contour of the spatially distributed chart object. Several hundreds of these complex accesses are usually necessary for the comparison of one radar image with the chart. Therefore, one of the accesses has to be accomplished within a few milliseconds.

The data structure of the database for the chart has to be object-oriented and flexible with respect to modifications. Object-orientation is necessary for the optimal interpretation and classification of the structures in a radar image. Flexibility is an obvious necessity for an experimental system, where changes due to new findings have to be feasible in a compatible and easy manner.

3.2 Objects in the electronic chart

Two different coordinate systems are used for the chart. One of them is an absolute coordinate system. For practical reasons, this system is the German official Gauss-Krüger system, although any similar system can be employed as well. The second coordinate system is used for the purpose of fast access to the data of the electronic chart. These so-called *river coordinates* employ the *river axis* as a reference line. The river axis is a virtual line along the river defined by the river authorities. It ideally consists of straight lines and curves of constant radius fitted together without bends. The position on the river is given by the kilometer along the axis and an offset perpendicular to the axis. This coordinate system is optimal for the purpose of access to objects in the neighborhood of a given point. The objects in the database for the electronic chart are sorted according to their longitudinal

	Real Objects	Virtual Objects
Spirals	River Bank	Ideal Guiding Line Limits of Navigable Water River Axis
Polygons	Bight Bridge Embankment Ferry Groyne Harbour Island Lock Mooring Overhead Line Point of Embarkation River Mouth	
Points	Buoy Landmark	Altitude Branch Level Station Mean Current Message River Identification

Table 1: Objects in the electronic chart

position in the river. Selecting an object with a known position only requires searching the database linearly, which can be implemented as a very fast operation. A two-dimensional search, on the other hand, would be time-consuming and hard to implement under the given real-time constraints.

The classes of objects held in the database are listed in table 1. Real and virtual objects are distinguished. Real objects are objects visible in the waterway such as the river banks, bridges, locks etc., whereas virtual objects express information relevant primarily for ship guidance, e.g. limits of the navigable water, ideal guiding lines, altitude.

The object classes in table 1 are also sorted according to their structure into objects for points, polygons and spirals. Point-like objects do not have an extent in the chart. Polygon objects represent structures in the chart with a limited extent, which can be depicted by one or more polygons with a bounded number of points. There is a one-to-one correspondence between point-like as well as polygon objects and a single entity in the chart.

Spiral objects, on the other hand, do not correspond to a single entity. An arbitrary number of these objects together describe a structure in the chart. This structure is oriented along the river, with every object equivalent to one interpolation point for the structure. Between two interpolation points the distance to the river axis is interpolated linearly, resulting in Archime-

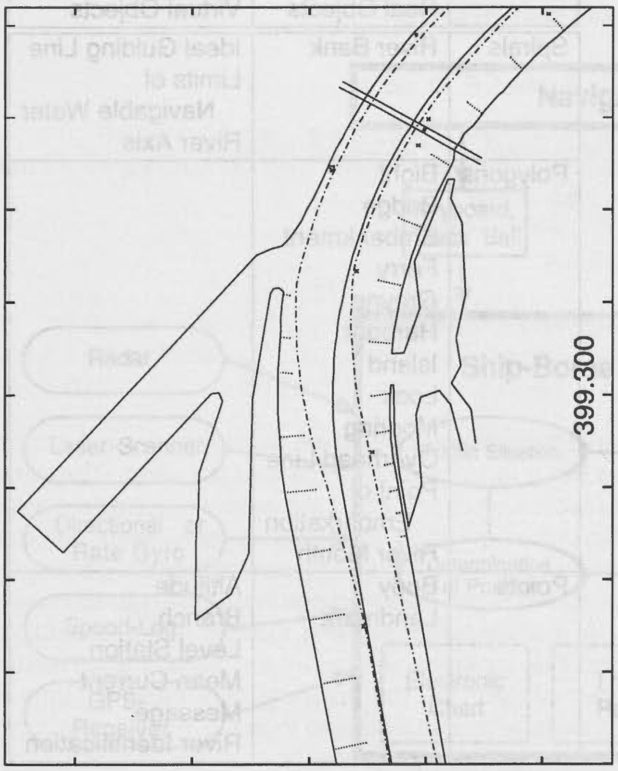


Figure 2: Example of an electronic chart

dian spirals. This type of interpolation is especially appropriate for a river and results in a small number of interpolation points.

The appearance of objects represented in the chart usually depends on the water level. This dependance is linearly interpolated in the chart by assigning a slope of the embankment to every spiral or polygon element. The value assigned can also be interpreted as a measure of accuracy of the element for image matching purposes.

An example of a hardcopy of the electronic chart display is shown in figure 2. Two harbors and a bridge can be easily distinguished. In the river, the river axis and the limits of the navigable water are plotted as solid respectively dashed-dotted lines.

More details about the implementation of the chart are given in [4].

4 MATCHING IMAGE DATA AND ELECTRONIC CHART

Within this section, a least squares technique for matching images and the electronic chart is explained. Input to the matching algorithm are predicted values of the ship's position and heading. Corrections to these predicted values are computed by minimizing a weighted sum of distances between image and electronic chart. These corrections and their accuracies are used as measurements in a Kalman filter described in section 6.

The matching technique is used for matching radar as well as laser scanner images with the electronic chart. Furthermore, stationary targets detected by a multiple-target tracking algorithm can be matched to landmarks in the electronic chart by the same technique.

4.1 Matching radar echo contours and electronic chart

The process of matching the radar contours and the electronic chart starts with an estimate of position and heading based on previous matching steps. For the current radar image, a prediction of the ship's position and heading has been computed according to the mathematical model of the own ship. Also actual measurements from other sensors like a gyroscope or GPS may have been processed by a Kalman filter algorithm, so this information is already integrated in the initial values of the matching process. The initial values consist of the global coordinates $s_{r(-)}$ and $s_{h(-)}$ of the position of the imaging sensor and the global heading $\psi_{(-)}$ of the ship. The acquisition of one radar image lasts some 2.3 seconds, the time for one complete turn of the radar antenna. As the ship is in general not stationary, the coordinates obtained by the matching process are assigned to the time when the antenna beam is pointing in the ship's heading direction. As mentioned before, the purpose of the matching process is the computation of corrections

$$\Delta s^* = \begin{bmatrix} \Delta s_r^* \\ \Delta s_h^* \\ \Delta \psi^* \end{bmatrix} \tag{1}$$

to the initial position and heading.

The standard radars used on commercial vessels generate a map-like image of the surroundings. The antenna beam has a horizontal beam width of 0.5 to 1.0 degrees, the vertical beam width is about 15 to 20 degrees. The antenna performs one complete turn within 2.3 seconds. During the turn, radar pulses of 50 ns length are sent out, the pulses are repeated with a maximum frequency of 3 kHz. Thus the raw radar image consists of about 7000 radial beams. The echo signals received by the antenna are digitized, filtered and transferred into the main memory of the computer. When filtering the raw image, a reduction of the number of rays is also performed. The radar data available in the computer are made up as a binary image consisting of 720 rays with 256 radial pixels. The radial quantization can be chosen as 3 or 6 m, thus resulting in a maximum sensing distance of 768 or 1536 m. Currently the interface to the radar is redesigned to enable a maximum number of 4096 radial pixels.

On the digitized, binary radar image some preprocessing steps are performed. The radar image is segmented into radar objects representing a continuous echo area. Also characteristic values like the area of the object and its extremal coordinates are determined.

Relevant echos for the matching process are gained from the radar data by selecting appropriate radar objects and extracting the front contour of these radar objects. Aggregating contour points located closely together, the echo points are deducted. Thus a further data reduction is achieved, resulting in about 200 echo points within one radar image. Figure 3 shows the measurement of one echo point before matching. The echo coordinates r_i and ϕ_i are given in a polar coordinate system assigned to the ship's body, the index i is used to distinguish the echo points within one radar image. As shown in figure 3 the nearest point on the contour of the electronic chart with its coordinates k_{r_i} and k_{h_i} is computed. The distance vector $\mathbf{d}_i$ is perpendicular to the chart contour. Based on the distance to the next chart contour a validation step is performed. If the distance exceeds a certain threshold, the assignment of echo and chart is assumed erroneous and the measurement is discarded. Erroneous assignments may arise from noise in the radar image or other vessels obviously not contained in the electronic chart. After the validation step n measurements are assumed to be available for further processing. In the determination of the nearest chart point, the motion of the ship is taken into account by using an individual position and heading s_{r_i}, s_{h_i}, ψ_i for each echo point. This is expressed by the following equation:

$$\begin{bmatrix} s_{r_i} \\ s_{h_i} \\ \psi_i \end{bmatrix} = \begin{bmatrix} s_{r(-)} \\ s_{h(-)} \\ \psi_{(-)} \end{bmatrix} + \begin{bmatrix} \Delta s_{r_i} \\ \Delta s_{h_i} \\ \Delta \psi_i \end{bmatrix} \quad (2)$$

The corrections Δs_{r_i}, Δs_{h_i} and $\Delta\psi_i$ are gained by interpolating and extrapolating the position estimates of the last radar image and the initial position of the matching process. This is done using the exact measuring time of echo i.

Based on figure 3 the following equation can be derived:

$$\begin{bmatrix} -r_i \sin(\varphi_i + \psi_i + \Delta\psi) \\ -r_i \cos(\varphi_i + \psi_i + \Delta\psi) \end{bmatrix} + \begin{bmatrix} s_{r_i} \\ s_{h_i} \end{bmatrix} + \begin{bmatrix} \Delta s_r \\ \Delta s_h \end{bmatrix} = \begin{bmatrix} k_{r_i} \\ k_{h_i} \end{bmatrix} \quad (3)$$

Δs_r, Δs_h and $\Delta\psi$ are the desired corrections to the initial position and heading. By linearizing these equations with respect to the variable $\Delta\psi$, taking into account equation (2) and neglecting higher order terms, equation (3) can be transformed to the following form:

$$\underbrace{\begin{bmatrix} 1 & 0 & e_{h_i} \\ 0 & 1 & -e_{r_i} \end{bmatrix}}_{\mathbf{M}_i} \underbrace{\begin{bmatrix} \Delta s_r \\ \Delta s_h \\ \Delta\psi \end{bmatrix}}_{\Delta \mathbf{s}} = \underbrace{\begin{bmatrix} d_{r_i} \\ d_{h_i} \end{bmatrix}}_{\mathbf{d}_i} . \quad (4)$$

All n measurements result in the system of equations:

$$\underbrace{\begin{bmatrix} \mathbf{M}_1 \\ \vdots \\ \mathbf{M}_n \end{bmatrix}}_{\mathbf{M}} \underbrace{\begin{bmatrix} \Delta s_r \\ \Delta s_h \\ \Delta\psi \end{bmatrix}}_{\Delta \mathbf{s}} = \underbrace{\begin{bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_n \end{bmatrix}}_{\mathbf{d}} . \quad (5)$$

The accuracy of matching is affected by many factors. Especially effects that impair the correspondence of radar image and electronic chart must be taken into account. Major impacts on the accuracy of matching are:

1. The accuracy of the radar sensor.
2. The nature of the specific radar objects and their influence on the reflection properties. The echo received for example from a concrete river bank is more reliable than one from a flat, sandy river bank.
3. The dependency of the echo locus on the actual water level.
4. The direction specific information of the distance vectors $\mathbf{d}_i$. They contain only information perpendicular to the chart contour.
5. Possible erroneous assignments of echos and electronic chart. Such assignments may be caused for example by ships located close to the chart contour.

The measurement accuracy of the radar sensor is a-priori known. It is assumed that the errors in the coordinates r_i and φ_i are independent and normally distributed with the variances σ_r^2 and σ_φ^2. In a local cartesian (ξ_i,η_i)-coordinate system according to figure 3 the following equation for the measurement errors $\Delta\xi_i$ and $\Delta\eta_i$ holds:

$$E\left[\begin{bmatrix} \Delta\xi_i \\ \Delta\eta_i \end{bmatrix}_l \begin{bmatrix} \Delta\xi_i & \Delta\eta_i \end{bmatrix}_m \right] = \begin{bmatrix} r_i^2\sigma_\varphi^2 & 0 \\ 0 & \sigma_r^2 \end{bmatrix} \quad (6)$$

The measurement errors belong to different measurements l and m in a (ξ_i,η_i)-coordinate system. The nature of the chart objects and their dependency on the water level can also be stored in the electronic chart. In the matching, the reliability of a echo i is

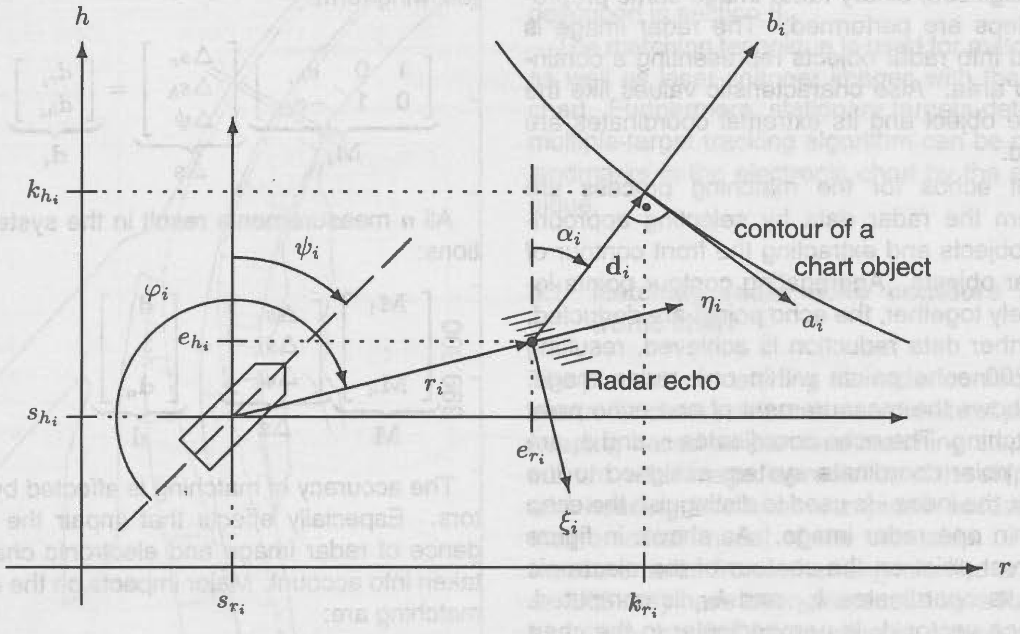


Figure 3: Geometric properties for matching

described by a weighting factor h_i in equation (7). So each measurement r_i, φ_i is weighted by a matrix given in the specific (ξ_i,η_i)-system by the equation:

$$\mathbf{W}_{\xi_i,\eta_i} = h_i \left[E \left[\left[\begin{array}{c} \Delta\xi_i \\ \Delta\eta_i \end{array} \right]_l \left[\begin{array}{cc} \Delta\xi_i & \Delta\eta_i \end{array} \right]_m \right] \right]^{-1}$$

$$= h_i \left[\begin{array}{cc} \frac{1}{r_i^2\sigma_\varphi^2} & 0 \\ 0 & \frac{1}{\sigma_r^2} \end{array} \right] \quad (7)$$

Furthermore the distance vectors contain information depending on their direction. As described, they are perpendicular to the chart contours and represent a distance information in this specific direction. As most of the chart objects are oriented along the waterway, most distance vectors are oriented transversal to the waterway. Thus there is only little information for matching in the longitudinal direction. The available information can be exploited appropriately, when the measurements are weighted vectorially. The weighting matrix $\mathbf{W}_{\xi_i,\eta_i}$ is transformed into the cartesian (a_i,b_i)-system. This system is oriented along the contour of the chart object. For matching, only the component in the b_i-direction is relevant. The weight in the a_i-direction is set to zero, as there is no information available. After a further transformation into the global coordinate system, the weighting matrix of a measurement is given by:

$$\mathbf{W}_i = \left[\begin{array}{cc} w_{11_i} & w_{12_i} \\ w_{12_i} & w_{22_i} \end{array} \right]$$

$$= \left[\begin{array}{cc} g_i \sin^2\alpha_i & g_i \sin\alpha_i \cos\alpha_i \\ g_i \sin\alpha_i \cos\alpha_i & g_i \cos^2\alpha_i \end{array} \right] \quad (8)$$

With:

$$g_i = \frac{h_i}{r_i^2\sigma_\varphi^2}\sin^2(\alpha_i - \varphi_i - \psi_i) + \frac{h_i}{\sigma_r^2}\cos^2(\alpha_i - \varphi_i - \psi_i). \quad (9)$$

With vectorial weighting the corrections for the initial position and heading can be computed as:

$$\Delta\mathbf{s}_R^* = \left[\mathbf{M}^T\mathbf{W}\mathbf{M}\right]^{-1}\mathbf{M}^T\mathbf{W}\mathbf{d} \quad (10)$$

$$\text{with} \quad \mathbf{W} = \left[\begin{array}{ccc} \mathbf{W}_1 & & 0 \\ & \ddots & \\ 0 & & \mathbf{W}_n \end{array} \right]$$

The result of matching radar image and electronic chart is the correction $\Delta\mathbf{s}_R^*$ to the initial position and heading and the covariance matrix $\mathbf{C}_R = \left[\mathbf{M}^T\mathbf{W}\mathbf{M}\right]^{-1}$ as a measure of the accuracy obtained by the matching process.

To illustrate the matching of radar echos and chart contours two radar images are given in figure 4 and 5. The own ship is displayed in the lower middle of the images as a small ship symbol. The grey areas are the radar echos acquired by the computer. The tick marks at the frame are 100 m apart. These images were recorded on the river Rhine near Mannheim. The electronic chart is superposed to the radar images. In figure 4 the distance vectors are also plotted. In this figure the initial values of the matching process have been changed manually to generate significant

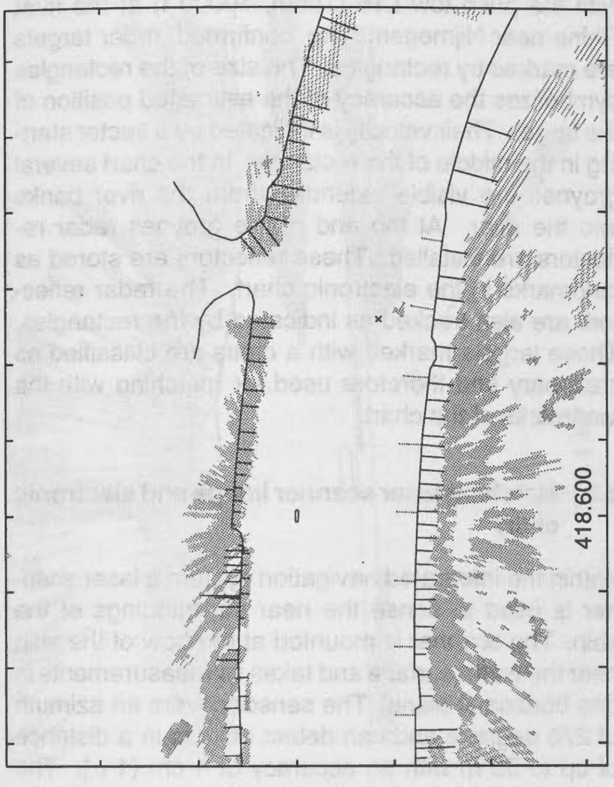


Figure 4: Distance vectors between chart and radar image

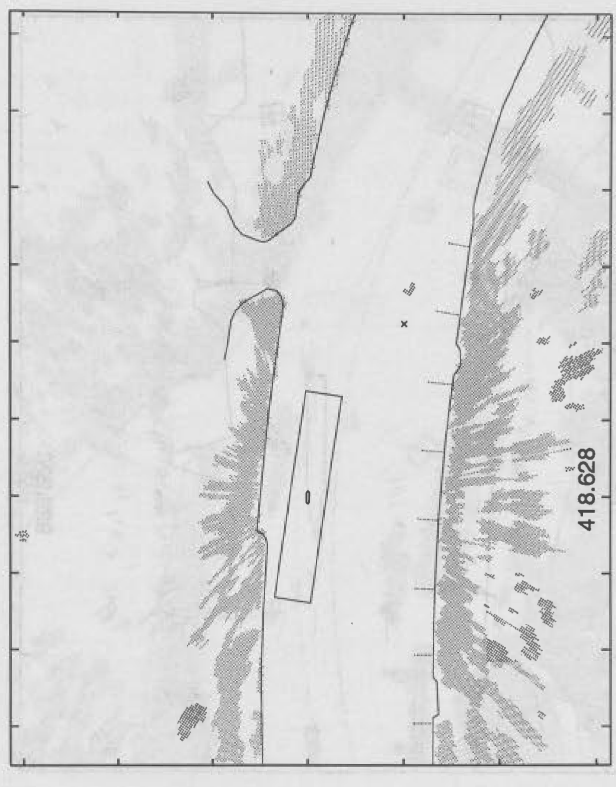


Figure 5: Accuracy of matching

distance vectors. In figure 5, the same radar image is shown. Now the electronic chart is superposed to the radar image based on the estimated position and heading of the vessel. Here the uncertainty of the matching process is visualized by the rectangle around the own ship. This rectangle symbolizes the 30 σ - confidence ellipsoid of the positional corrections. It is evident that the accuracy of matching in transversal direction is much higher than in the longitudinal direction.

4.2 Matching stationary targets and electronic chart

The matching algorithm described in section 4.1 provides enough information about the position and heading of the ship on most inland waterways. Its basic requirement is that the major part of the contours stored in the electronic chart are also visible in the radar image. This is always true for canals with almost constant water level. On a free flowing stream like the major part of the river Rhine, the water level can vary widely. At the lower Rhine area from Duisburg towards the North Sea, large parts of the river banks are often flooded and therefore no longer visible. To allow a safe navigation on that part of the river, the banks and the groynes are equipped with radar reflectors that remain visible even at a high water level. These radar reflectors are stored as landmarks in the electronic chart. They are used by another matching process that will be described in this section.

A multiple-target tracking algorithm (cf. section 5) is implemented within the integrated navigation system. This algorithm searches for radar objects on the waterway and tracks them through an image sequence. When a radar object is found that cannot be associated with an existing track, a new track is initiated. As there has been no other information on this track, it is marked as 'tentative'. If there have been measurements for this new track in some subsequent radar images, the estimates of the position and velocity of the tracked object become more accurate and the track is marked as 'confirmed', meaning that the track is with a high probability caused by a really existing object.

Based on the estimated velocity and its covariance the tracked objects can be classified as stationary or moving. For the purpose of matching with the electronic chart only confirmed tracks also classified as stationary are used. The track coordinates estimated by the multiple-target tracking algorithm are used as image measurements. In a next step, the distance vectors from the image measurements to the nearest landmarks of the electronic chart are computed. The motion of the own vessel during the acquisition of the radar image is already taken into account by the

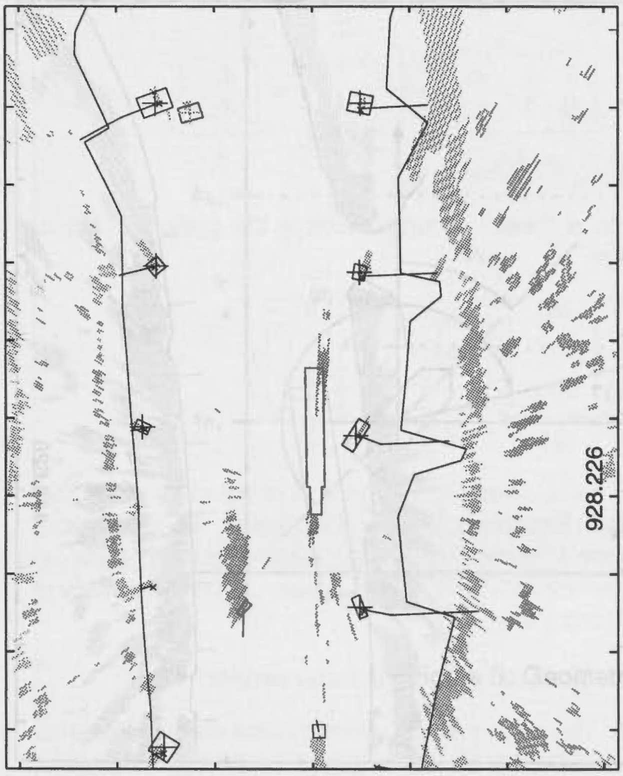


Figure 6: Example for matching stationary targets and chart

tracking algorithm. Thus a correction of the position and heading of the own ship according to equation (2) is not needed.

Based on the distance vectors a validation step is performed. All assignments with a distance larger than a certain threshold are assumed erroneous and the image measurements are discarded. Another condition can be stated to avoid wrong associations. A landmark of the chart can only be assigned to one stationary object. If there is more than one validated assignment, only the assignment with the smallest distance is taken into account for matching purposes. The distance vector is defined by the difference of positions of two assigned points and therefore contains two dimensional displacement information. Thus the weighting of the measurements must be different from the algorithm described in section 4.1. In this case, the inverse of the position covariance matrix of the specific tracking filter can be used for weighting the measurements. With the computed distance vectors and the modified weighting matrices the corrections are computed as discussed in section 4.1.

The result of matching stationary targets and electronic chart is the correction $\Delta \mathbf{s}_S^*$ and the covariance matrix $\mathbf{C}_S = \left[\mathbf{M}^T\mathbf{W}\mathbf{M}\right]^{-1}$ as a measure of the accuracy obtained by the matching process.

An example of this matching procedure is displayed in figure 6. The radar image was recorded at a trial with the push tow L16 (185m, 10000 t) at the river Rhine near Nijmegen. The 'confirmed' radar targets are marked by rectangles. The size of the rectangles symbolizes the accuracy of the estimated position of the target. Their velocity is indicated by a vector starting in the middle of the rectangle. In the chart several groynes are visible extending from the river banks into the river. At the end of the groynes radar reflectors are installed. These reflectors are stored as landmarks in the electronic chart. The radar reflectors are also tracked as indicated by the rectangles. Those targets marked with a cross are classified as stationary and therefore used for matching with the landmarks of the chart.

4.3 Matching laser scanner image and electronic chart

Within the integrated navigation system a laser scanner is used to sense the near surroundings of the ship. The scanner is mounted at the bow of the ship near the water surface and takes its measurements in one horizontal plane. The sensor covers an azimuth of 270 degrees and can detect objects in a distance of up to 50 m with an accuracy of 4 cm (1 σ). The radar sensor has considerable disadvantages in detecting the near surroundings of the ship. The least detectable distance is about 15 m and the radial quantization of the digitized radar image is 3 m. Thus laser scanner measurements are essential for the navigation in narrow canals, for entering locks or for docking manoevers.

Although the coverage and the accuracy of laser scanner and radar are quite different, the images derived from both sensors are similar, because both sensors yield a map-like image of the surroundings. Thus the same matching techniques as explained for the radar in section 4.1 can be used for the laser scanner image. The implementation of the matching algorithm for laser scanner images is subject to current work within this project.

Figure 7 shows an example of the laser scanner image. This picture was recorded when entering a lock chamber on the river Neckar. The black points are the laser scanner measurements, the grey line is the contour of the lock from the electronic chart.

5 TRACKING OBJECTS IN IMAGE SEQUENCES

Besides time invariant objects, e.g. river banks, the navigation environment scanned by the radar sensor also comprises foreign ships, anchored buoys and radar reflectors. In order to obtain information about the actual traffic situation, it is essential to reconstruct the trajectories of the latter objects, with respect to position, speed and orientation in relation to the own ship [7]. A bank of dynamic object models has to be processed in parallel, each model describing an

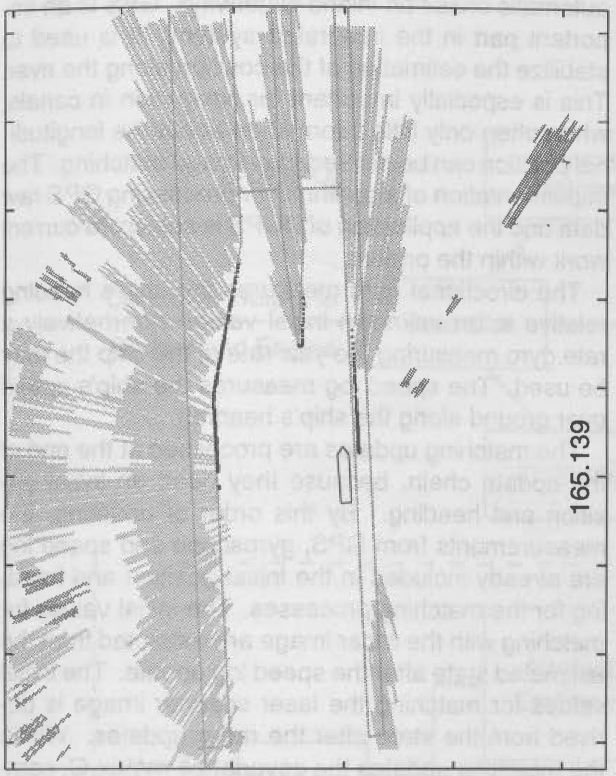


Figure 7: Example of a laser scanner image

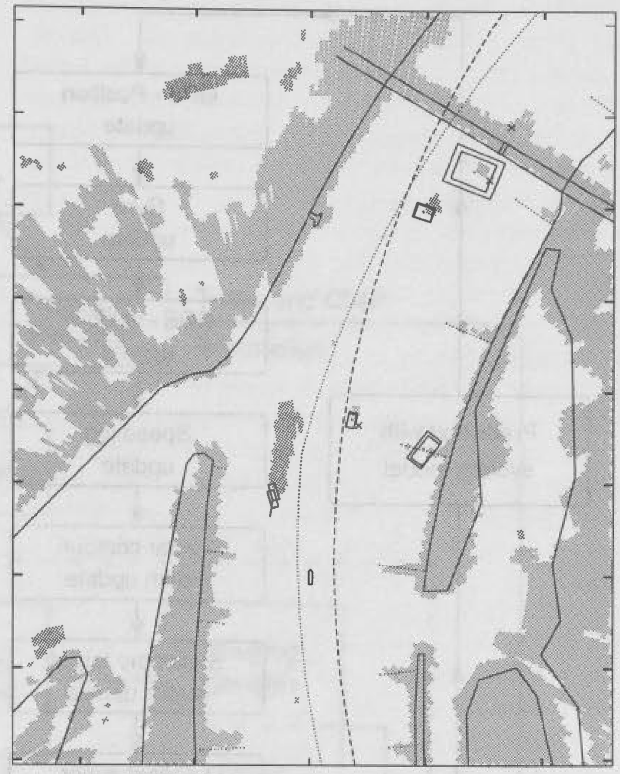


Figure 8: Example for multiple-target tracking methods

assumed target track. This kind of information can not be gained from single radar images. It is necessary to interpret the chronological development in image sequences. This goal is achieved by a multiple-target tracking algorithm [1]. Methods applied are Kalman filtering [2] for the estimation of target trajectories given the data-association history, and recursive Bayesian tests [8] in order to sequentially interpret the arriving radar image measurements. This results in a set of contradicting hypotheses, each of which is associated with a certain probability. The basic problem originates from the exponential growth of hypotheses arising if all possible measurement associations are considered. A detailed description of the implemented algorithms can be found in [6].

Figure 8 gives an example of a real radar image, showing one oncoming ship and several radar buoys. Here we can distinguish two gates for each track, corresponding to the update and prediction steps of each Kalman filter. The line originating at the center of the prediction gate symbolizes the estimated speed vector. From this image it can be seen, that by means of the target tracking technique, a reliable reconstruction of the navigation scenery can be obtained.

Applying the tracking algorithm to inland shipping, one has to deal with radar echos arising for example from bushes or trees at the river banks or from bridges. Therefore a classification step is included in the image processing, where relevant objects for tracking are selected. Within this step, the information of the electronic chart is used. The distances from a radar object to the nearest objects in the chart provide additional classification information. So it is possible to decide, whether a radar object is outside the river and thus irrelevant for tracking. The detection of radar buoys and landmarks and the treatment of echos arising from bridges or overhead lines can also be improved by using the chart information. Thus the application of chart information improves the efficiency of the tracking algorithm.

6 INTEGRATION OF MEASUREMENTS

The integration of results obtained from different matching algorithms and the measurements of other sensors is performed by an extended Kalman filter [2]. Within the filter, the different measurements are weighted according to their actual accuracy and an integrated estimate of the ship's state is computed. Designing the filter, one has to take into account that the accuracy of the measurements may vary with time. Measurements may even be not available temporarily. When for example no landmarks are visible, the matching of stationary targets and landmarks cannot be performed. Also some sensors may not be installed on a specific ship.

In addition to the measurements derived from the

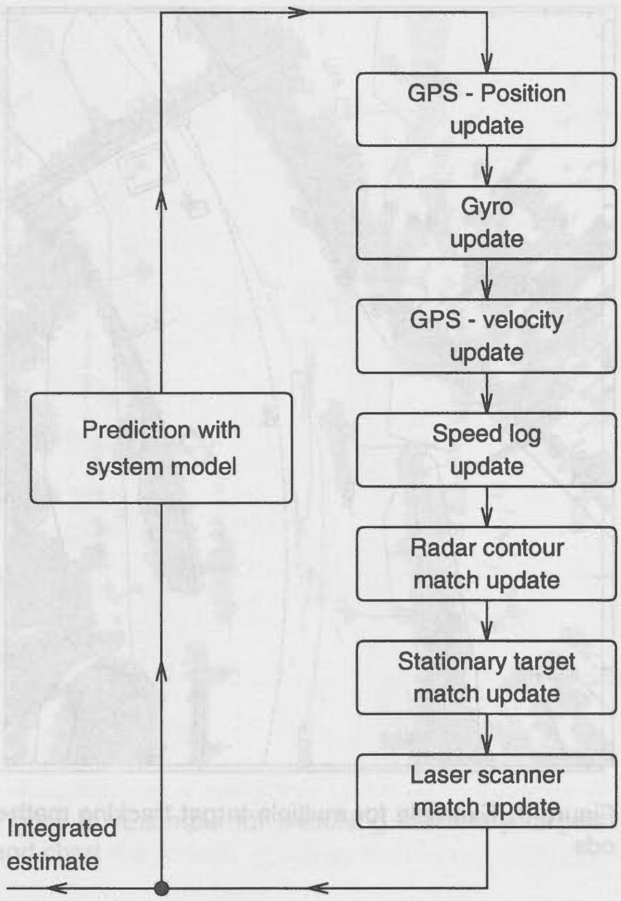


Figure 9: Flow of information within the Kalman filter

imaging sensors, the extended Kalman filter also processes the measurements from a GPS receiver, a ultrasonic speed log and a directional gyro.

Figure 9 shows the processing steps within one Kalman filter cycle. The filter uses a nonlinear system model with 9 states to perform the prediction of the ship's state and its covariance. A detailed discussion of the system model can be found in [5]. The measurements of the different sensors are processed in separate update steps. This makes it easy to account for not available measurements by passing state and covariance unchanged through an update block.

As GPS measurements, the position and velocity computed by the GPS receiver itself are presently used. They are transformed into global chart coordinates and used as separate updates. This is due to the experience, that the GPS velocity obtained from the receiver is more reliable than the position. GPS measurements are not continuously available to the navigation system. The reception of the satellite signals is interrupted when the ship passes a bridge. Also satellites may be not visible when cruising in narrow valleys or near to trees on a river bank. Although the accuracy obtained solely from GPS in the standard positioning service is not sufficient for an automatic cruise on inland waterways, GPS is an important part in the integrated system. It is used to stabilize the estimation of the position along the river. This is especially important for navigation in canals, when often only little information about the longitudinal position can be derived from image matching. The implementation of algorithms for processing GPS raw data and the application of DGPS is subject to current work within the project.

The directional gyro measures the ship's heading relative to an unknown initial value. Alternatively a rate gyro measuring the yaw rate of the ship the may be used. The speed log measures the ship's speed over ground along the ship's heading.

The matching updates are processed at the end of the update chain, because they need an initial position and heading. By this order of updating, the measurements from GPS, gyroscope and speed log are already included in the initial position and heading for the matching processes. The initial values for matching with the radar image are extracted from the estimated state after the speed log update. The initial values for matching the laser scanner image is derived from the state after the radar updates. Within the matching updates the covariance matrix $\mathbf{C}$, computed in the matching process, is used as the covariance of the measurement noise. Thus the filter can account for the actual accuracies obtained from matching.

The integrated estimation of the ship's state is given by the state vector after the laser scanner update.

7 GENERALIZED STRUCTURE OF THE INTEGRATED NAVIGATION SYSTEM

The generalized structure of the integrated navigation system is displayed in figure 10. The integrated navigation system consists of a sensor processing level (dashed box in figure 10) and a control level for optimization and control.

Generally the information obtained directly from the sensors is not sufficient for guiding the ship. As a consequence, a-priori knowledge of the ship's dynamics and the features of the navigation environment must be contributed to the sensor information. The upper part of the sensor processing level represents the real world, the ship, its navigation environment and all sensors available on the ship. The model world in the lower part of this diagram is a reproduction of reality in the computer by means of the available a-priori knowledge consisting of dynamic models for the own ship and other vessels and the electronic chart of the waterway. To match the real world with the model world the sensor signals are used. Difference signals are deduced by comparing the sensor signals to equivalent signals generated from the model world. Examples for such difference signals are the corrections to the initial values of the image matching

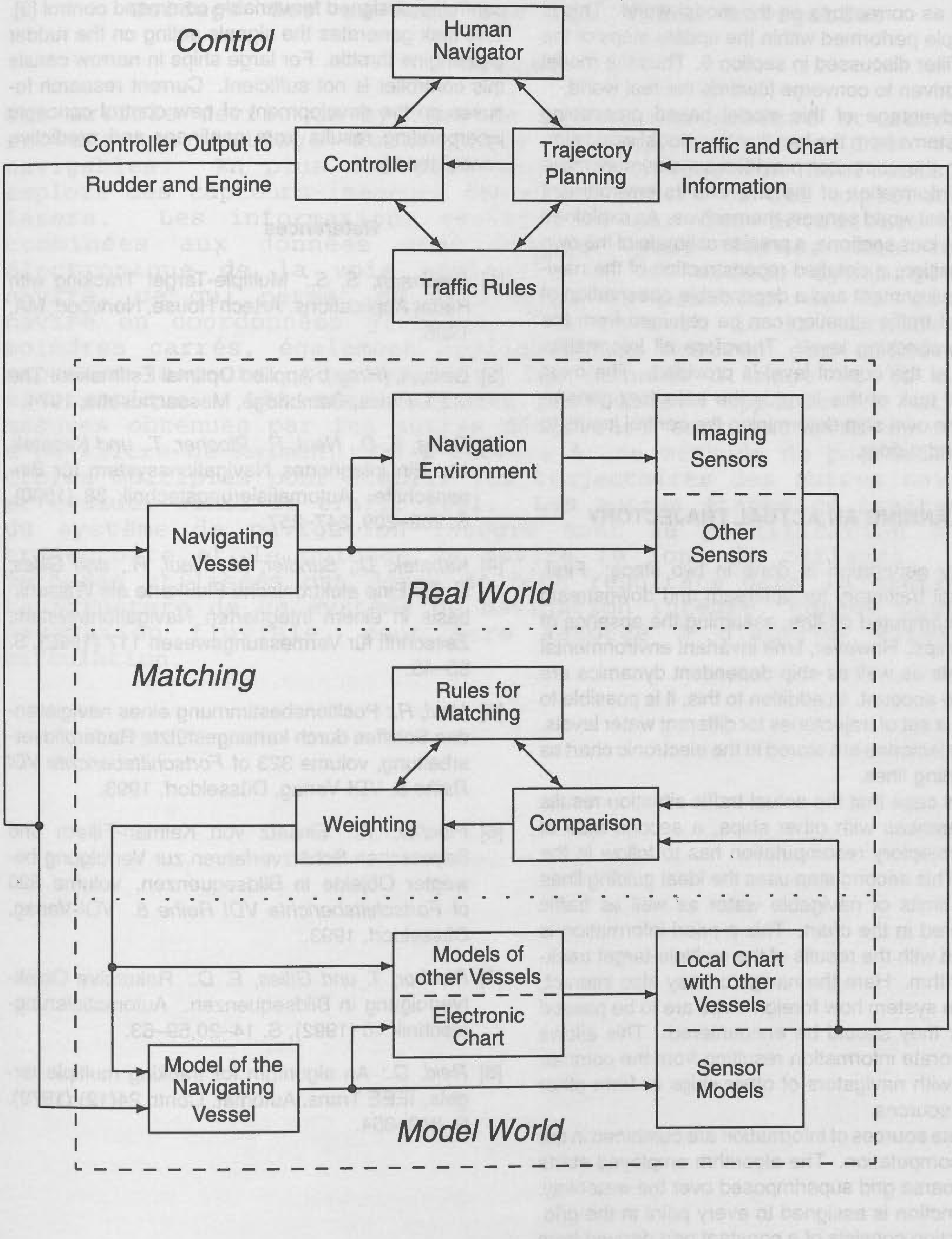


Figure 10: Structure of the integrated navigation system

algorithms. The matching algorithms themselves are part of the block *comparison* in figure 10. The difference signals are weighted according to the actual situation and the accuracy of the measurements and are used as corrections on the model world. This is for example performed within the update steps of the Kalman filter discussed in section 6. Thus the model world is driven to converge towards the real world.

The advantage of this model based processing scheme stems from the fact that the model world simulated on the computer provides considerably more detailed information of the ship and its environment than the real world sensors themselves. As explained in the previous sections, a precise estimate of the own ship's position, a detailed reconstruction of the navigation environment and a dependable observation of the actual traffic situation can be obtained from the sensor processing level. Therefore all information required at the control level is provided. The most important task of this level is the trajectory generation for the own ship determining the control inputs to engine and rudder.

8 PLANNING AN ACTUAL TRAJECTORY

Trajectory generation is done in two steps: First, an optimal trajectory for upstream and downstream travel is computed off-line, assuming the absence of foreign ships. However, time invariant environmental constraints as well as ship dependent dynamics are taken into account. In addition to this, it is possible to compute a set of trajectories for different water levels. These trajectories are stored in the electronic chart as ideal guiding lines.

For the case that the actual traffic situation results in interferences with other ships, a second step of on-line trajectory recomputation has to follow in the sequel. This second step uses the ideal guiding lines and the limits of navigable water as well as traffic rules stored in the chart. This a-priori information is combined with the results of the multiple-target tracking algorithm. Here the navigator may also interact, telling the system how foreign ships are to be passed and how they should be encountered. This allows to incorporate information resulting from the communication with navigators of other ships or from other external sources.

All these sources of information are combined in the on-line computation. The algorithm employed starts from a coarse grid superimposed over the waterway. A risk function is assigned to every point in the grid. This function consists of a constant part derived from the chart information and a time-varying part representing the results of the multiple-target tracking algorithm. A foreign ship is taken into account at a predicted place of encounter. The second step in the on-line calculations is a search algorithm resulting in possibly several trajectories through the grid and a cumulative risk for each trajectory. Finally, one of these trajectories is selected as input for the control task.

The control task is implemented as a linear state controller designed for variable command control [3]. This task generates the signals acting on the rudder and engine throttle. For large ships in narrow canals this controller is not sufficient. Current research focuses on the development of new control concepts incorporating results from nonlinear and predictive control theory.

References

[1] *Blackman, S. S.:* Multiple-Target Tracking with Radar Applications. Artech House, Norwood, MA, 1986.

[2] *Gelb, A. (Hrsg.):* Applied Optimal Estimation. The M.I.T. Press, Cambridge, Massachusetts, 1974.

[3] *Gilles, E. D., Neul, R., Plocher, T., und Kabatek, U.:* Ein integriertes Navigationssystem für Binnenschiffe. Automatisierungstechnik 38 (1990), S. 202–209, 247–257.

[4] *Kabatek, U., Sandler, M., Neul, R., und Gilles, E. D.:* Eine elektronische Flußkarte als Wissensbasis in einem integrierten Navigationssystem. Zeitschrift für Vermessungswesen 117 (1992), S. 35–45.

[5] *Neul, R.:* Positionsbestimmung eines navigierenden Schiffes durch kartengestützte Radarbildverarbeitung, volume 323 of *Fortschittsberichte VDI Reihe 8.* VDI-Verlag, Düsseldorf, 1993.

[6] *Plocher, T.:* Einsatz von Kalman-Filtern und Bayesschen Schätzverfahren zur Verfolgung bewegter Objekte in Bildsequenzen, volume 320 of *Fortschittsberichte VDI Reihe 8.* VDI-Verlag, Düsseldorf, 1993.

[7] *Plocher, T. und Gilles, E. D.:* Rekursive Objektverfolgung in Bildsequenzen. Automatisierungstechnik 40 (1992), S. 14–20,59–63.

[8] *Reid, D.:* An algorithm for tracking multiple targets. IEEE Trans. Automat. Contr. 24(12) (1979), S. 843–854.

Guidage des navires sur les voies navigables

Résumé

Cet article décrit un système de navigation intégré et autonome servant au guidage automatique des navires sur les voies navigables. En plus des détecteurs de navigation habituels, on emploie des capteurs-imageurs comme des radars ou des explorateurs lasers. Les informations recueillies par les détecteurs sont combinées aux données déjà connues pour former une carte électronique de la voie navigable et un modèle dynamique du navire, ce qui permet de déterminer la position et le cap du navire en coordonnées globales. On se sert de la méthode des moindres carrés, également appliquée aux mesures des différents capteurs-imageurs, pour apparier les données d'image et la carte électronique. Les résultats des processus d'appariement et les mesures obtenues par les autres détecteurs sont intégrés au moyen d'un filtre de Kalman. On a recours à une méthode de poursuite de cibles multiples pour établir les trajectoires des autres navires et évaluer ainsi le trafic réel. Les autres étapes de traitement du système de navigation intégré sont la planification d'une trajectoire et le guidage du navire le long de celle-ci. Les méthodes utilisées ont donné de bons résultats dans la pratique. La structure de ce système de navigation intégré, conçue pour la navigation fluviale, peut être adaptée à d'autres voies de circulation.

AIRBORNE SENSOR SYSTEMS

Prof. Dr.-Ing. Gunther Schanzer

Dipl.-Ing. Detlef Kayser
Institute for Flight Guidance
Technical University of Braunschweig
GERMANY
Tel.: ( ); Fax: ( )

ABSTRACT

In the present time the field of navigation is undergoing a revolutionary development. There are three major technology aspects that lead to this leap in performance of todays navigation systems:

- the introduction of satellite navigation (GPS/GLONASS)
- the development of a new generation of inertial sensors (e.g. sensors on the chip)
- the increased computer capabilities that allow complex real time calculations

Especially in airborne applications the new technologies will lead to systems that can achieve centimetre accuracy in all three dimensions, in all phases of flight at very low prices compared to todays avionic equipment.

The Institute for Flight Guidance from the Technical University of Braunschweig operates a twin engined research aircraft that is equipped with a satellite based navigation system. The paper will present the basic structure of this integrated navigation system and give examples of various applications of the system and show the achieved performance. Special emphasis will be given to the least results in long range differential (over 400km range) experiments and the possibilities to use GLONASS, the Russian counterpart to GPS. The Institute for Flight Guidance is also doing research on new, low cost inertial sensor packages that are based on gyro on the chip technology. By error modelling and inflight calibration techniques it is planned to significantly improve the performance of this sensor class. The test philosophy and first test results will be presented in the paper.

Groupes-capteurs aéroportés

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif

Upgrading of Stereoplotters with New Hardware and Software Components

Klaus Szangolies
Special Project Leader

Closewitzer Strasse 44
D - 07743 Jena/ Germany

KEY WORDS: Photogrammetry, Instruments, Optical, Hardware, Software, Superimposition

ABSTRACT

The activities for upgrading of the widely used stereoplotters have to be welcomed and deserve support, because they increase the efficiency of photo-grammetry in the transition phase from analog and analytical technology to digital technique.

KURZFASSUNG

Die Modernisierung und Erweiterung der zahlreichen genutzten Stereoauswertegeräte verdient Unterstützung, weil damit die Effektivität der Photogrammetrie in der Phase des Übergangs von der analogen und analytischen Technik zur digitalen erhöht und gesichert wird.

1. Introduction

A poll held within the framework of the special project "Upgrading Photogrammetric Instruments" (1992 to 1994) ascertained the opinions of equipment manufacturers and service companies with regard to the modernization/complementation of photogrammetric plotting instruments. More than ten explained their strategies.

The paper further considers a very informative topical article from "Geodetical Info Magazine" (GIM), No. 12, of December 1993 having the title "GIM Interviews the Photogrammetry Industry" [6].

The opinions ascertained differ very much.

These are the two most extreme positions:

- The transition from stereoplotting with analog, opto-mechanical plotting instruments (the instruments still most widely used) via analytical plotting to the predominant use of digi-tal photogrammetry will still take ten to twenty years.
- In 1996 already, the upgrading of analog and analytical plotters will be meaningless, because "digital photogrammetry will dominate world-wide".

The reality certainly lies anywhere between these two positions. And since according to cautious estimates approximately 5000 analog plotters and several hundred analytical plotters are currently used in pratice, we have to consider how to further proceed with them.

Without doubt, the availability of new equipment, its efficiency and price/performance ratio will be decisive criteria as to whether
- available equipment will be upgraded, or
- new equipment will be purchased and the old one rejected.

It is necessary to distinguish between
◊ the modernization of existing equipment, e. g. by the exchange of obsolete computers and software for new ones, or the replacement of mechanically connected drawing tables by electronically controlled plotters, and
◊ the complementation of existing equip-ment, e. g. the extension of the viewing system of stereoplotters by superimposition or a higher viewing magnification.

2. Subject of upgrading

2.1. Analog, opto-mechanical stereo-plotters

This category of instruments, which were produced approximately between 1965 and 1985, includes among others:

Stereometrograph,Topocart: Zeiss-Jena
Planimat, Planicart:Zeiss- Oberkochen
Autograph A7, A8, A10, B8, Aviomap: Wild
PG 2, PG 3: Kern
G6, G7, G8: Galileo
IIC, III, IV:Santoni
SFP, SDF:OMI

Typical complementations/extensions made on these instruments:

- Digital output of the X, Y and Z coordinates for transmission to a coordinate recording instrument, a PC or an electronically controlled digital drawing table
- Digital output of the photo coordinates x', y', x", y" for transmission to a coordinate recording instrument or a PC (e. g. for analytical aerotriangulation)
- Digital input of the orientation data φ, ω, κ, bx, by, bz (received from aerotriangulation or computer-assisted orientation) into a plotter
- Conversion of analog into analytical plotters
- Expansion of the optical system to allow simultaneous observation of photo and map (superimposition)
- Increase of the viewing magnification, e. g. in the case of the PG 2 from 2X - 4X - 8X (standard) to 3X - 6X - 12X or 4X - 8X - 16X.

2.2 Analytical plotters

The older instruments of this category, produced approximately between 1980 and 1985, include among others:

Planicomp C 100:Zeiss Oberkochen
Aviolyt AC 1, BC 1, BC 2: Wild
DSR 1:Kern
APC 3, 4, 5: OMI
Digicart, Stereocart:Galileo
Traster:Matra
SD-4:QASCO

Typical complementations/ modernizations:

- Exchange of obsolete electronic control equipment
- Exchange of computers meanwhile obsolete: Data General Nova, Hewlett Packard, DEC, etc. for advanced PCs with 386 or 486 processors based on MS/DOS or UNIX, or for Microstation
- Exchange of software meanwhile obsolete (fittingly described by a renowned company as "fossilware") for modern one
- Linkage of modernized analog or analytical plotters to comprehensive plotting systems such as PHOCUS (Zeiss), INFOCAM (Leica), etc.

2.3 Orthoprojectors

The most widely used orthophoto instruments include:

Orthocomp Z 2: Zeiss
Avioplan OR 1: Wild.

Upgrading is made similarly to that of analytical plotters (Section 2.2).

3. Summary

By upgrading, the efficiency of existing older photogrammetric plotting instruments can be increased and their life extended by five to ten years. A number of renowned and newly founded companies offer excellent services in this area (see Section 4).

Digital photogrammetric equipment is more and more gaining ground. In the year 2000, analog and analytical plotters will, however, not completely be replaced by it for a long time yet [6].

There are company representatives advancing the opinion that in 1996, when the next ISPRS Congress will be held, the upgrading of the old instruments will no longer be necessary and will be of academic significance only. They believe digital photogrammetry will have fully replaced conventional photogrammetry at this time.

This opinion seems to be unrealistic. What can be expected is that the transition to digital photogrammetry will be steadily continued in map production and in the generation of GIS data banks. It will, however, still take ten years or more until it is completed. And during this time we need well functioning plotting instruments as well as techniques mastered by their operators.

4. References

[1] ADAM Technology, PE&RS 1993,p.795
[2] APY Photogrammetric Systems GmbH, PE&RS 1993,p.795
[3] Carl Zeiss, Brochures 1993
[4] DAT/EM Systems International, Brochures 1993
[5] DAT/EM, Tomorrow's Digital Mapping Technology, PE&RS 1993, p.807/808
[6] GIM Interviews the photogrammetric Industry. Geodetic Info Magazine- GIM, Lemmer/ The Netherlands, 12/1993,p.31-34
[7] International Systemap Corp. ISM Brochures 1993
[8] Leica Inc. PE&RS 993, p.835/836
[9] QASCO Analytical Systems, GIM 5/1993,p.22
[10] QASCO, New Upgrade for AP's, PE&RS 1993,p.1207

[11] Schwebel,R., Upgrading of Stereoplotters by Carl Zeiss, Presented Paper ISPRS-Symposium Com.II, Ottawa, June 1994

[12] Stapley,K., Analytical Stereoplotters - Another Midlife Crisis? ITC- Symposium 21.April 93 ISPRS-Symposium Com.II, Ottawa, June 1994

[13] Wong,P., Upgrading Photogrammetric Instruments. Presented Paper ISPRS-Symposium Com.II, Ottawa, June 1994

[14] Yates,J.F., Upgrading Photogrammetric Instruments. Presented Paper ISPRS-Symposium Com.II, Ottawa, June 1994

Amélioration des appareils de restitution photogrammétrique au moyen de nouveaux matériels et logiciels

RÉSUMÉ

MOTS CLÉS : PHOTOGRAMMÉTRIE, INSTRUMENTS, OPTIQUE, MATÉRIEL, LOGICIEL, SURIMPRESSION

RÉSUMÉ :

Il convient d'encourager et d'appuyer les activités visant à améliorer les stéréorestituteurs d'emploi courant car ces travaux permettent d'accroître l'efficacité de la photogrammétrie dans la phase de passage de la technologie analogique et analytique à la technologie numérique.

Upgrading Photogrammetric Instruments

Patrick Wong, MA

I.S.M. International Systemap Corp.

Suite 777 - 650 Georgia Street, Vancouver, B.C. V6B 4N7

Tel:(604) 684-3750 Fax:(604) 684-9750

ABSTRACT

Analog photogrammetric instruments can be upgraded by encoding/interfacing to CAD software. Encoding system designs include the utilization of rotary or linear encoders mounted on the photo or model coordinate axes. Typical interfacing technology utilize solid-state "cards" for popular PC-Compatible CPU's. CAD software system may have been created specially for digital mapping or adapted from generic off-the-shelf packages.

A more radical approach is the full conversion to analytical operations. Such approach requires special mechanical engineering and computerized "plate control" system. A better integration with CAD software is also effected.

The cost/benefit analysis however, may not justify the upgrading of a typical analog instrument any longer due to the availability of affordable digital image based systems. In any case, the demand for digital image products can never be supported by upgraded analog or analytical instruments.

KEY WORDS

Instrument encoding, calibration, digital mapping system, stereo superimposition.

1.0 INTRODUCTION

Photogrammetric instruments based on the simulation (or analog) principle have been in production for more than 45 years. Many of the early models of these, so called analog instruments, are still in use world-wide despite the advent of the analytical plotter and digital image workstations. Instruments that are still in active duties include WILD A7, A8, A10, AMU/H, B8, AG1; ZEISS D2/3, E2/3; JENA Stereometrograph & Topocart models; Kern PG2 and some Santoni models.

In the late 1970's it was apparent that Analytical Plotters and digital mapping were to be the way of the future. Many manufacturers offered retrofit systems for the existing late model instruments so that digital mapping could be performed. All of these "up-grades" included some form of instrument encoding with an interface device connected to a computerized graphic system. For example, WILD offered the Synercom-Infomap interface, ZEISS offered the Intergraph IGDS interface and KERN offered the PC-PRO interface.

By early 1980's however, the major manufacturers had all but withdrawn from up-grading analog instrument and were concentrating their efforts on Analytical Plotters. Despite the lack of support from the major manufacturers, a thriving after-market instrument up-grading industry developed. I.S.M. International Systemap Corp. (ISM), amongst others, offers such services.

One approach to up-grading is a full "analytical" conversion of the instrument. This approach is more radical in that the process is irreversible. Most of these designs retain nothing more than the viewing and the plate carrier sub-systems of the original instrument. The "Space Rod" and other mechanism that were used to maintain a parallax-free model are replaced with a servo-motor driven system controlled by analytical photogrammetry software.

A more conservative approach in analog instrument up-grading may include instrument encoding, viewing systems modification and graphic system interfacing. ISM concentrates on providing a "turn-key" service in this approach.

This paper will report ISM's experience in the field and provide some recommendations based on those experiences. This paper will also offer some of our insights as to the future of analog instrument upgrading.

2.0 GENERAL CONSIDERATIONS IN INSTRUMENT UP-GRADING

ISM realized from the outset that digital systems are "unforgiving". Imperfections in the instruments (instrument errors) are brutally revealed in numerical form. It is therefore extremely important that Users should understand that an instrument in good repair is necessary to support a digital conversion. ISM has also learned that many instruments were installed in less than ideal environments with little or no maintenance. In order to avoid confusing instrument errors with software algorithm problems, ISM always offered, and occassionally insisted on a full instrument service and calibration performed along with the up-grading.

Many late model analog instruments were factory encoded because the instruments were intended for "digital" aerial triangulation or to be connected to a semi-automatic electrically driven plotting table instead of the mechanical coordinatograph. Instruments of this type include:

Table 1

Supplier	Instrument Model	Plotting Table
WILD	A10, AMH/U, AG1	TA, TA2, TA10
ZEISS	D3, E3	Datatech
KERN	PG2	Datatech

The factory encoding systems installed were mechanically sound, however the encoder resolutions were intended for hard copy map plotting and were usually not fine enough for typical digital mapping applications. Often, the encoder output were non-TTL or the encoders would require a special power supply. In either case, the casing for the factory encoders were large enough that fitting a replacement was never a problem. The factory encoder resolution may be determined by the encoder model numbers. ISM maintains a reference list for all popular encoders.

ISM normally recommends that the encoders be replaced with a modern TTL/+5volt unit so that power may be supplied directly from the computer. The recommended resolution is 10um in model space. Finer resolution can, of course be effected, however the encoders would most likely be measuring the "noise" or the imperfection of the mechanical systems.

Most digital mapping system entails that the map be entirely compiled and edited in the digital domain. Hard copy, if required, is then plotted with a pen or ink-jet plotter driven directly by the computer. The co-ordinatograph or plotting table attached to the instrument is now redundant.

Many photogrammetrists however, still feel uncomfortable without a plotting table directly driven by the instrument. Often, they would request that the digital up-grade be configured in such a way that the plotting table and the computer system operate independently but simultaneously. This type of installation often induces electronic interference problems that are difficult to isolate and resolve, however a switching mechanism can be inserted into the system so that either digital or mechanical plotting may be performed conveniently.

In addition to instrument conditions, site conditions are also very important. It cannot be overly stressed that clean and uninterrupted electrical power be maintained.

3.0 ENCODING SYSTEMS

There are many designs for encoding systems. In general, they can be classified broadly as photo coordinate systems and model coordinate systems. The mechanical principles are essentially simple. A photo coordinate based system will measure the positions of the left and right viewing microscopes (x'y';x"y") as they traverse the photo space. A model coordinate based system will measure the position of the measuring mark (x,y,z) as it traverses the model space. Of course, the interfacing software must also be able to accept either the photo or model coordinate inputs in order to effect either system.

3.1 ENCODERS

Modern TTL encoders are available in the Rotary or Linear types. A Rotary encoder utilizes a glass disc with fine markings subdividing a complete rotation (360 degrees) into equal sectors. If mounted concentric to the lead-screw, the encoder will then subdivide each complete lead-screw rotation (pitch) into fine equal parts.

eg.:

If the model stage lead-screw pitch is 2mm(2000um),
then a 250 count/rev. encoder will yield a model space resolution of 2mm/250 or 8um.

A Linear Encoder is essentially a glass scale with very fine markings. 5 or 10um spacing per marking (least count) is generally sufficient. A Linear Encoder can be used to provide continuous encoder counts instead of a rack-and-pinion/rotary encoder system.

A modern TTL encoder is normally rated at 50,000 hours life or more, however if malfunctioning occurs, it is cheaper to replace the entire device.

3.2 MOUNTING SYSTEMS

The mounting of encoders to the instrument can be problematic and costly. A good mounting system will provide easy and accurate mounting with the least disturbance to the instrument. Some of the design considerations discussed are:

a. Accuracy

The encoders are expected to measure the fine motions of either the model stage or the viewing microscopes. If a Linear Encoder is used, it is important that it is mounted parallel to the guide-rail or lead-screw. If a Rotary Encoder is used, then the axis of encoder rotation should be parallel to the rotary shaft or lead-screw.

To minimize time wasted in adjusting or calibrating the errors, ISM has developed various mounting jigs for most instrument types. The technician simply attached the jig onto certain critical instrument component (eg. a main guide rail) and drill and tap screw holes for attaching the entire encoder mounting bracket set. This method ensures that the encoder mounting is within the precision tolerance and easy to install.

b. Standardization

ISM has selected as few encoder types and resolutions as possible for the instrument types encountered. This enables bulk purchasing of the devices and maintaining adequate stock.

For the mounting bracket systems, many of the component are interchangeable between instrument types. Sizes of machine bolts and other fasteners are also minimized and standardized. Simple hand tools are all that is required to install a system.

c. Disturbance to the instrument

All systems with the exception of the B8/B8S and Jena SMG encoding systems will not disturb the calibration of the instrument. Some systems are so simple that a User may self-self-install with only simple hand tools.

Encoding systems have been developed for the following instrument types:

Table 2

Instruments	Encoding	Comments
WILD A7	3-Axis R/L	supports free motion
WILD A7	3-Axis R	no free motion measurement
WILD A8	3-Axis R/L	supports free motion
WILD A8	3-Axis R	no free motion measurement
WILD B8/B8S	4-Axis L	
WILD A10	3-Axis R/L	supports free motion
WILD AG1	3-Axis L	
ZEISS D2/3	3-Axis R/L	
ZEISS E2/3	3-Axis R/L	
ZEISS F2/3	4-Axis L	
Jena Topocarts	3-Axis R 4-Axis L	
Jena SMG's	3-Axis R	
SANTONI G6	3-Axis R	
3-Axis		encoding in the model space x,y,z.
4-Axis		encoding in the photo space x'y';x"y".
R		encoding with rotary encoder
L		encoding with linear encoder

For hand-wheel instruments with free motion capabilities such as WILD A7's, A8's & A10's, two versions of encoding are available. The simpler version (3-Axis R) is mounted on the instrument gear box and does not support free motion while a more advanced version utilize linear encoders mounted along the model stage lead-screw providing for continuous encoder count whether the instrument is in free motion mode or not.

Most of the encoding mounting hardware utilize existing holes already drilled in the instrument chassis. Sometime holes exposed after removal of a sub-component in the instrument are utilized.

3.3 WIRING HARNESS

ISM makes extensive use of modern plastic wiring harnesses available for the computer or telecommunication industry. For example, standard RJ-11 telephone jacks are use to connect encoder leads to the main cables so that positive connections and easy removal are achieved. This will also eliminate soldering of wiring which is prone to breakage if handled carelessly. The ultimate objective is total system reliability.

The wiring harness gathers the multitudes of fine wiring into a convenient bundle terminating in a common connector such as RS232 or D50. The 50 pin D50 is most commonly used because it cannot be confused with any connectors or ports found in a modern computer.

4.0 INTERFACE DEVICES

Interface devices have been available from major manufacturers for some time, however these devices were typically designed for coordinate data recording in the mensuration process for aerial triangulation. For example, WILD offered EK-5, and later EK-22, which were commonly delivered with the A7, A8 & A10 models. Such devices can convert the signals from the encoders into co-ordinate values (typically model co-ordinates) for aerial triangulation programs.

These devices are bulky, troublesome and by-and-large obsolete. The modern equivalent is a simple solid state digitizing "card" for the computer CPU.

Digitizing cards were not originally designed for photogrammetric instrument up-grading. They can be found in other digital applications as part of the controlling system. For example, numerically controlled machine tools such as a milling machine might utilize a 3 channel card to control the "longitudinal feed", "cross feed" and "milling depth" of the cutting tool. Such a 3 channel card may also be used to accept model co-ordinate (x,y,z) inputs.

In the course of time various computers ranging from "mini's", Workstations and PC have been used for instrument upgrading, however this field has been dominated by Personal Computers ever since the introduction of the Intel 80286 based systems. Manufacturers such as Altec have been supplying digitizing cards for some time, but the most popular card is the "SEC-PC" Card manufactured by Fischer Computer Systems for the PC-AT bus. This card is used by almost all vendors in this field.

The function of a digitizing card is to monitor the status of the encoders and ancillary devices (such as Foot Switches) in the instrument. Should there be a change in the position of the microscopes (4-Axis) or the model stage (3-Axis), the digitizing card will register the new position(s) and send the new data to the interface software system.

Digitizing cards are extremely reliable and requires no User maintenance. Of the hundreds of systems ISM has delivered, only 2 cards had been returned as damaged. In one case the cause was a severe lightning strike to the building which had also damaged several computers. The other was found to be short circuit due to User mishandling. In any case, should a malfunction ever occur, replacement can be delivered readily via air courier without undue delay.

5.0 INTERFACE SOFTWARE & GRAPHIC SYSTEM

The interface software for a digital up-grade typically includes 4 major components:

- a facility for the orientation of the stereo model,
- a facility for numeric data output,
- a facility for graphic data output
- a graphic system.

In the ISM SysteMap Digital Mapping Software, the orientation module accepts either photo or model coordinate input. In the model coordinate mode, it provides computer assisted Absolute Orientation. In the photo co-ordinate mode, it provides Interior Orientation facility as well. In both cases, the User must first perform Relative Orientation. This facility establishes the Matrix which is used to transform the encoder signals (x',y';x",y" or x,y,z) in real-time into ground coordinates (X,Y,Z) for the numeric or graphic data output facilities.

The facility for numeric data output may support tasks such as mensuration for aerial triangulation. The system may output photo coordinates (only in the 4-Axis case and in stereo comparator mode), model coordinates (after Relative Orientation has been performed) and ground coordinates (after Absolute Orientation has been completed).

The facility for graphic data output is essentially a "driver" for the graphic system. The Driver directs and inserts ground coordinates (X,Y,Z) into the graphic system which generates map symbology that constitute a digital "manuscript".

Graphics systems that have been used in Analog Instrument upgrading include Intergraph IGDS, Synercom, Atlas, Kork, AutoCAD, Intergraph MicroStation, CadMap and others. ISM SysteMap utilizes MicroStation exclusively.

6.0 DIGITAL CALIBRATION

ISM provides digital calibration of up-graded instruments. The methodology published by Prof. Klaus Szangolies (1966) is utilized. The calibration data is recorded directly from the encoders and a 15 point (3 x 5) stereo grid test is performed.

ISM utilizes an in-house software system to indicate the adjustments necessary, to record the observations, to perform the calculations, to generate the model deformation diagram and to generate the horizontal and vertical accuracy ratings of the model.

This kind of calibration is particularly important because a modern digital map is essentially a "positional" file of mapped features. The Positional Files together with the Digital Elevation Model (DEM) are the data base from which cartographic representations (eg. contour map) are derived. A structured Positional File is also a primary source of data for a Geographic Information System (GIS). The positional accuracy, which is entirely dependent on the instrument accuracy, cannot be over-stressed.

7.0 FINANCIAL CONSIDERATIONS

A complete budget for instrument up-grading must include the following elements:

a. Instrument cost

The capital cost of the instrument may justifiably be overlooked by some businesses with long term ownership of instruments. The cost of instruments has often been fully recovered or completely depreciated within the financial structure of the business, however if an instrument has to be acquired, then the combined cost of 2nd-hand purchase, packing/removal, shipping, installation/overhaul and calibration cannot be overlooked.

b. Encoding cost

Even if the encoding process has been greatly simplified by ISM, a significant cost still remains. If the instrument has been factory encoded, the cost is far less as replacement encoders are relatively inexpensive.

c. Computer cost

Fortunately, Mini Computers or Workstations are no longer mandatory in instrument up-grading. Personal Computers are the dominant choice. A basic PC may be fairly cheap. However, a system with good CPU speed, RAM, adequate Hard Disk Drive, and large monitor is still a cost consideration.

d. Software cost

A complete digital mapping software system will comprise of an instrument interface system, a graphic system, TIN/CIP programs for the DEM and ancillary system for data integrity checking/repairs. The total cost of a complete system is not inexpensive.

e. Installation and training cost

A major budget item is the fees and expenses payable for the installation of the instrument/encoding system and computer hardware/software system. The fees payable on software training can also be substantial.

Often, A User will presume that self-training is possible, however this assumption would be false economy especially for a commercial installation. The time and business opportunity wasted will often cost more than the money saved.

8.0 CONCLUSIONS

An up-graded analog instrument, can produce a digital map file indistinguishable from that produced by a modern analytical instrument. Given proper instrument calibration, the accuracy of some late model 1st-Order instruments can even be compared favourably to an analytical instrument.

From a productivity point of view, an up-graded analog instrument can be superior to an analytical instrument because of wider choice of graphic and data structuring/checking software from third-party sources. The limited automation provided by an analytical instrument in model set-up is more than overcome by superior compilation software systems.

Since the introduction of ISM DiAP (Digital image Analytical Plotter) in the 1992 ISPRS Congress in Washington, a full up-grade may no longer be advisable. Often, the total cost of acquiring and up-grading an instrument is equal to that of a complete DiAP system.

Additional benefit offered by DiAP such as real-time stereo colour super-imposition of the digital compilation or digital orthophoto option are not available nor technically feasible for an analog instrument at all. In addition, the lack of spare parts (some of them critical) and the maintenance cost of an analog instrument must also be seriously considered.

REFERENCES

Szangolies, Klaus Von, 1966. Vorschläge zur einheitlichen Testung und Bewertung von Stereoauswertegeräten, Eingegangen: 1. 3. 1966 pp. 156-189

Perfectionnement des instruments photogrammétriques

RÉSUMÉ

On peut perfectionner les appareils analogiques de photogrammétrie par codage ou interfaçage avec des logiciels CAO. La conception des systèmes de codage comprend l'utilisation de codeurs rotatifs ou linéaires montés suivant les axes de coordonnées d'une photographie ou d'un modèle. Pour l'interfaçage, on a habituellement recours aux cartes à circuits intégrés conçues pour l'UCT des micro-ordinateurs compatibles et largement répandus. Les systèmes informatiques de CAO peuvent avoir été créés spécialement pour la cartographie numérique ou provenir de l'adaptation d'ensembles matériel/logiciel standards.

Le passage complet à l'exploitation analytique constitue une approche plus radicale qui exige des dispositifs mécaniques particuliers et un système informatisé de «contrôle d'épreuves». Une meilleure intégration aux logiciels CAO est également mise en application.

Cependant, il est possible que l'analyse coûts-avantages ne justifie plus l'amélioration des instruments analogiques typiques, vu la disponibilité de systèmes abordables basés sur des images numériques. De toute façon, l'utilisation d'instruments analogiques ou analytiques améliorés ne pourrait satisfaire la demande pour des produits sous forme d'images numériques.

Upgrading of Stereoplotters by Carl Zeiss

Dr. Reiner Schwebel
Carl Zeiss, Oberkochen

Abstract

Carl Zeiss offers upgrade packages to enable the integration of analog stereoplotters into state-of-the-art plotting systems (PHOCUS, CADMAP, MicroStation, AutoCAD). The Planicomp analytical plotters can be enhanced to the latest performance level through computer and software upgrades. Of special interest is the possibility of converting the C100 Planicomp into P2 Planicomp. The sturdy design of the viewers and the comprehensive service and support ensure both continuity and progress for the Carl Zeiss photogrammetric systems.

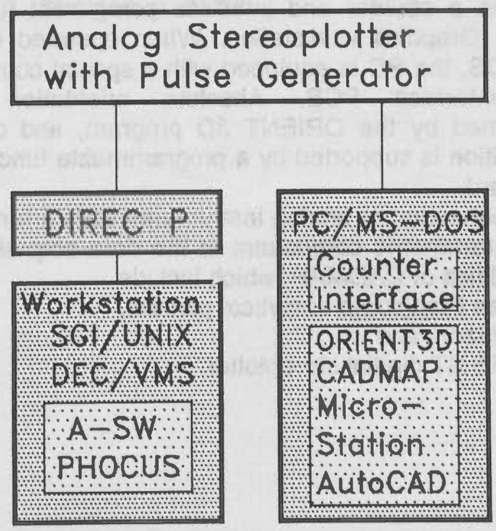
Key words: Digital mapping, stereoplotter upgrades

1. Introduction

Until the late seventies, the worldwide success of photogrammetry in the production of maps and plans was based on the development and industrial manufacturing of analog stereoplotters. Typical examples are A8, B8, PG2, Stereometrograph, Planimat and Planicart. Even in the age of analytical photogrammetry, however, the rugged design of these instruments, the service provided by the manufacturers and the upgrading options available guarantee a place for analog stereoplotters in the production of photogrammetric data.
The development of the early analytical plotters, such as C100 Planicomp, dates back to the seventies. The modular design of this instrument category, comprising an opto-mechanical viewer, electronics, computer and software, permits the user to replace individual components by advanced versions and to benefit from progress especially in the field of computers and software. Again, it has been demonstrated that opto-mechanical viewers like that of the C100 Planicomp may last for a service life of 20 years and more, provided that their design and construction are sufficiently sturdy.
Carl Zeiss as the leading manufacturer of photogrammetric systems therefore sees it as its responsibility to make upgrade packages available to the users of analog instruments and analytical plotters to enable the integration of photogrammetric viewers into a state-of-the-art production environment.

2. Upgrading of analog instruments

The XYZ coordinate movements of the model carriage in analog instruments have to be digitized and fed to a computer. For this, the analog instruments are fitted with linear or rotary encoders. The pulses are collected in a counter and interface unit and transmitted to the computer. After absolute orientation, the model coordinates are transformed in real-time into the national coordinate system. The plotting software processes the transformed coordinates into geometric object data, which is filed as a digitial map or GIS record and can be output.
Depending on the plotting software and computer platform used, Carl Zeiss offers three different upgrading options:



2.1 Integration into PHOCUS

The integration of analog instruments into PHOCUS, the photogrammetric cartographic system from Carl Zeiss, is achieved by means of

- the DIREC P counter and interface unit on a PC basis

and the

- A-software for absolute model orientation.

This permits all data acquisition and editing functions of PHOCUS - insofar as they can be applied in analog instruments - to be used for 3D data acquisition. In line with the PHOCUS structure, the geometric data is collected with the correct topology and is stored together with the attribute data in an object-oriented form in the PHOCUS data base. The data acquisition process is supported by the PHOCUS panel, a programmable function keyboard.

PHOCUS and the A-software run on a workstation platform: either on Silicon Graphics computers under UNIX or on DEC VAX or DEC AXP computers under VMS. The software can be implemented on its own computer to create a PA workstation, or it can be integrated as a PA workplace into the PHOCUS multi-user environment. The connection of analog instruments adds a further data source to the data acquisition options of PHOCUS, which comprise
P-Series Planicomp analytical plotters,
2D digitizers,
PHODIS ST digitial stereplotter and
monoplotting with the PHOCUS PM workstation.

2.2 Plotting with CADMAP

CADMAP is a photogrammetric CAD system from Carl Zeiss for digital mapping and GIS data acquisition, featuring particularly efficient acquisition and editing functions. CADMAP runs on the UNIX and MS-DOS platform.

For the operation with UNIX, the DIREC P is again used as a counter and interface component for a Silicon Graphics workstation. When operated with MS-DOS, the PC is equipped with a special counter and interface PCB. Absolute orientation is performed by the ORIENT 3D program, and data acquisition is supported by a programmable function keyboard.

The connection of analog instruments adds a further photogrammetric component to the data acquisition capabilities of CADMAP, which include
P-Series Planicomp analytical plotters,
2D digitizers and
PHODIS ST digitial stereplotter.

2.3 Plotting with MicroStation and AutoCAD

Carl Zeiss offers hardware and software packages for the integration of analog instruments into the MicroStation and AutoCAD CAD systems under MS-DOS. The package comprises:

- Counter and interface modules integrated into the PC
- ORIENT 3D program for absolute orientation and
- CADMAP/dgn MicroStation driver for MicroStation 5 or
- CADMAP/dwg AutoCAD driver for AutoCAD 12
- Function keyboard

The driver permits the acquisition and processing of 3D coordinates in a MicroStation or AutoCAD environment. The data is stored in the dgn or dwg format and output e.g. in the DXF format.

The CAD working environment has been enhanced by special photogrammetric CADMAP functions which are not available in the CAD packages. This includes, for example, the digitization of lines using tube increments, the continuous display of 3D coordinates and the cutting, extension, bending and closing of lines.

The commands are entered via standard and user-defined menues on the monitor or via a function keyboard.

3. Upgrading of analytical plotters

Analytical plotters comprising an opto-mechanical viewer, control and interface electronics, the host computer and the operating and plotting software allow the upgrading of each individual module to the latest standard. Upgrading is mainly centered on the electronics, the computer and software.

3.1 Upgrading of C100 Planicomp to P2 Planicomp

The C100 Planicomp launched by Carl Zeiss in 1976 and supplied until 1987 is operated by the HP 1000 real-time computers, while the LOOP real-time computation and control are handled by the host computer.

With the introduction of the P-Series Planicomp (P1, P2, P3 Planicomp), a microprocessor (P-processor) integrated into the instrument electronics is used for real-time computation. The P2 Planicomp is based on the time-tested viewer of the C100 equipped with the P-processor and a standard interface (RS232, IEEE-488) with the host computer.

An upgrade package comprising the

- P2 electronic system, work surface and software

permits the conversion of a

- C100 Planicomp with HP 1000 and Planimap software

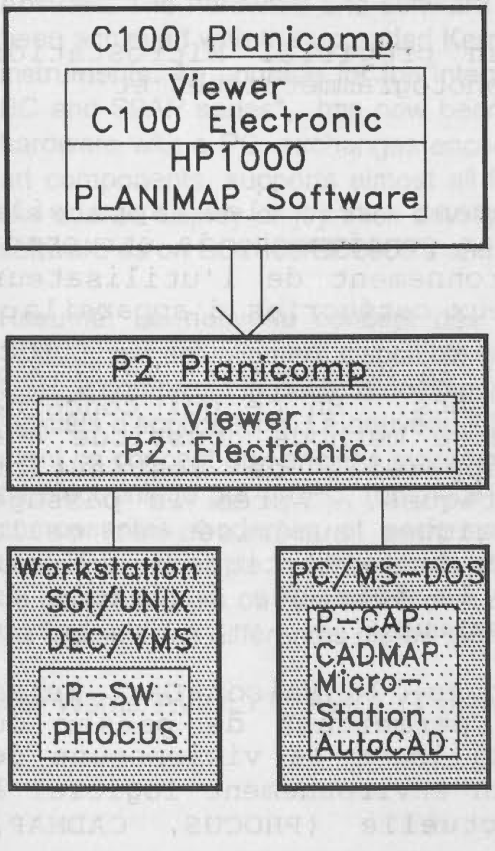
into a

- P2 Planicomp with P-processor and plotting software.

Depending on the computer platform, the converted P2 Planicomp can be run with

- P-software and PHOCUS on a UNIX or VMS basis,
- P-software and CADMAP on a UNIX basis, or
- P-CAP and CADMAP, MicroStation or AutoCAD on a MS-DOS basis.

This means that the user is able to operate the original C100 viewer in a state-of-the-art working and computer environment



3.2 Upgrading of P-Series Planicomp

The P-Series Planicomp was initially based on the host computers HP 1000 A (RTE A) and DEC MicroVAX (VMS). The availability of PHOCUS on Silicon Graphics workstations (UNIX) and DEC AXP (VMS) now permits the previous computers to be replaced by high-performance computers and earlier software versions by the new PHOCUS revisions 5 and 6. In MS-DOS configurations, new powerful PCs and the new revisions of the CAD software such as MicroStation 5 or AutoCAD 12 can now be used. In addition, the viewers of P3 and P33 Planicomp can be equipped with VIDEOMAP 30, the superimposition system for MS-DOS, which further increases their efficiency.

4. Practical implementation

Counselling and the implementation of the upgrading are the responsibility of the specialist personnel at Carl Zeiss Oberkochen and at the regional service and support centers. Advice, trainging and support are provided by photogrammetry experts, whereas the installation and integration of the new system components are performed by Zeiss service technicians and are frequently combined with a maintenance service for the opto-mechanical viewers.

References

Hobbie, D., 1987.
Indroduction to the new generation of Zeiss Products: Planicomp P-Series/PHOCUS Proceedings 41st Photogrammetric Week, Stuttgart, p. 21-24

Saile, J., 1987:
Performance data of the Planicomp P-Series.
Proceedings 41st Photogrammetric Week, Stuttgart, p. 25-28.

Modernisation des stéréorestituteurs par la Maison Carl Zeiss

RÉSUMÉ

La Maison Carl Zeiss offre toute une gamme de matériels et de logiciels permettant de moderniser les stéréorestituteurs analogiques et analytiques.

Des appareils analogiques numérisés connectés à des ordinateurs (OP, station de travail) permettent l'orientation absolue et la restitution par ordinateur. Le processus de restitution est fondé sur les systèmes suivants :

. système cartographique-photogrammétrique PHOCUS (station de saisie à entraînement de papier, UNIX ou VMS),

. système de CAO photogrammétrique CADMAP (MS-DOS, UNIX),

. CADMAP/dgn avec MicroStation, un progiciel MicroStation (MS-DOS) étendu par des fonctions photogrammétriques et

. AutoCAD (MS-DOS).

Tous ces progiciels tournent également avec les appareils analytiques de la série Planicomp P. Par conséquent, la structure des données, les fonctions et l'environnement de l'utilisateur sont pratiquement identiques pour les deux catégories d'appareils.

Cette modernisation a été appliquée par la Maison Carl Zeiss aux appareils analogiques de divers fabricants. En ce qui concerne les fonctions et la durée de vie, cette méthode constitue une solution de transition pour les appareils de restitution photogrammétrique numériques et analytiques. Après le passage entre l'utilisation d'appareils analogiques numérisés et celle d'appareils de restitution numériques ou analytiques, il est possible de continuer de se servir à la fois des logiciels de restitution et de ceux propres à l'ordinateur.

Dans le domaine des appareils analytiques, la Maison Carl Zeiss offre un progiciel de modernisation permettant de passer du Planicomp C100 au Planicomp P2, en utilisant la visionneuse de haute qualité du Planicomp C100 dans un environnement logiciel à la fine pointe de la technique actuelle (PHOCUS, CADMAP, MicroStation, AutoCAD).

En permettant de moderniser les appareils de restitution photogrammétrique, la Maison Carl Zeiss offre à la fois le progrès et la continuité dans le domaine de la fabrication des appareils photogrammétriques.

Upgrading the AVIOLYT Family of Leica's Analytical Instruments to the Leica Photogrammetric Workstation

Alfons Meid

Leica AG PMU, 5035 Unterentfelden, Switzerland

Symposium COM II , Ottawa 1994

Abstract: The hardware and software concept of the Photogrammetric Workstation has been achieved with the upgraded Kern DSR instruments and the Leica SD2000/SD3000 instruments. An upgrade for the integration of the AVIOLYT family, consisting of WILD BC and S9AP series*, has now been developed. This upgrade replaces the computer hardware with a PC, exchanges encoders and servo control electronics by state-of-the-art components, supports almost all features that are typical for the AVIOLYT family - like coarse display or joy stick driving of the stages - and allows for running the same software as on SD2000/SD3000's and upgraded DSR's.

Résumé: Le nouveau concept des Stations Photogrammétriques Leica SD2000 et SD3000 est actuellement disponible en kit de conversion pour tous les systèmes KERN DSR. Il comprend la partie processeur en temps réel sur PC et la gamme complète des logiciels d'applications. Le nouveau kit de conversion AVIOLYT permet d'adapter également ce concept aux anciens systèmes Wild BC et S9AP* ! Ce kit échange toute la partie électronique par un PC, remplace les codeurs linéaires et les servomoteurs par des composantes modernes et performantes. Il supporte et améliore les éléments de contrôle typiques de la série des Aviolyts tels le déplacement rapide des clichés à l'aide de commutateurs ou l'utilisation d'un manche à balai (joy stick) et permet d'utiliser tous les logiciels des différentes plates-formes de la série des SD2000/SD3000.

Key Words: AVIOLYT, AVIOLYT Upgrade

* An AC1 upgrade is currently evaluated. Le kit de conversion pour Wild AC1 est actuellement à l'étude.

Perfectionnement de la famille AVIOLYT d'appareils analytiques de Leica en station de travail photogrammétrique Leica

1. Introduction

The purpose of this paper is to explain the integration of WILD AVIOLYT analytical photogrammetric instruments into the concept of the Photogrammetric Workstation by upgrading their hardware and software. Chapter 2. gives a brief description of the Workstation concept and refers to previous publications on that topic. Chapter 3. explains all details of the upgrade.

2. The Leica Photogrammetric Workstation

The Leica Photogrammetric workstation is designed to provide complete analytical solutions for 3D data capturing. Its hardware and software concept enable users to perform all necessary steps to obtain 3D data in various CAD or GIS systems. This concept is now implemented on the WILD AVIOLYT family of analytical instruments, without removing their typical features.

2.1 Hardware Components

The Leica Photogrammetric Workstation Concept supports various analytical instruments (Refer to Fig. 1):

- the stereo digitisers SD2000 and SD3000 (in the following referred to as SD);
- the upgraded *Kern* DSR instruments (in the following referred to as DSR);
- the upgraded BC line of WILD AVIOLYT instruments (referred to as BC). Only the BC will be covered in this paper because the upgrades of the other WILD analytical instruments are not yet definite: an S9-AP upgrade is possible as well, but this might not be interesting for most of the S9 users. And the possibility of an AC1 upgrade is not yet sure (Feb. 1994).

All instrument types are controlled from a DOS PC, the Leica Mapping Terminal (LMT). Various host platforms may be connected to the LMT by serial communication to allow users to work on applications they prefer. Leica offers GIS/Mapping application software on PC-DOS, PC-Unix and VAX-VMS platforms. Moreover, there are other companies that provide application software connected to the LMT.

The SD line can optionally be equiped with the raster graphics based superimposition system COLORISS, available in mono and stereo.

A superimposition system for the DSR instruments, which is conceptually identical to COLORISS (but probably monochrome), is under development. Since the concept of LMT integrates the superimposition independently from the type of instrument, the same superimposition system may be integrated into the BC environment. However, the implementation has not yet been decided (Feb. 1994). It will depend on customer's request.

2.2 The LMT Software Concept

The central unit of the LMT is the real time program (RTP[1]), which is the same program for all instruments. It handles the real time loop operations, provides a general interface to host application software for 3D digitising and a user interface for an easy change of real time loop parameters or hardware settings, e.g. the speed of the stages or the illumination of photographs on SD instruments.

For more details about the concept see (*Cogan, Hinsken 1992*) and (*Hinsken, Meid 1993*); for detais focusing on close range photogrammetry and superimposition see (*Hinsken, Meid 1993*).

---

1 the AVIOLYT Upgrade is running with version 2.0 of RTP or later

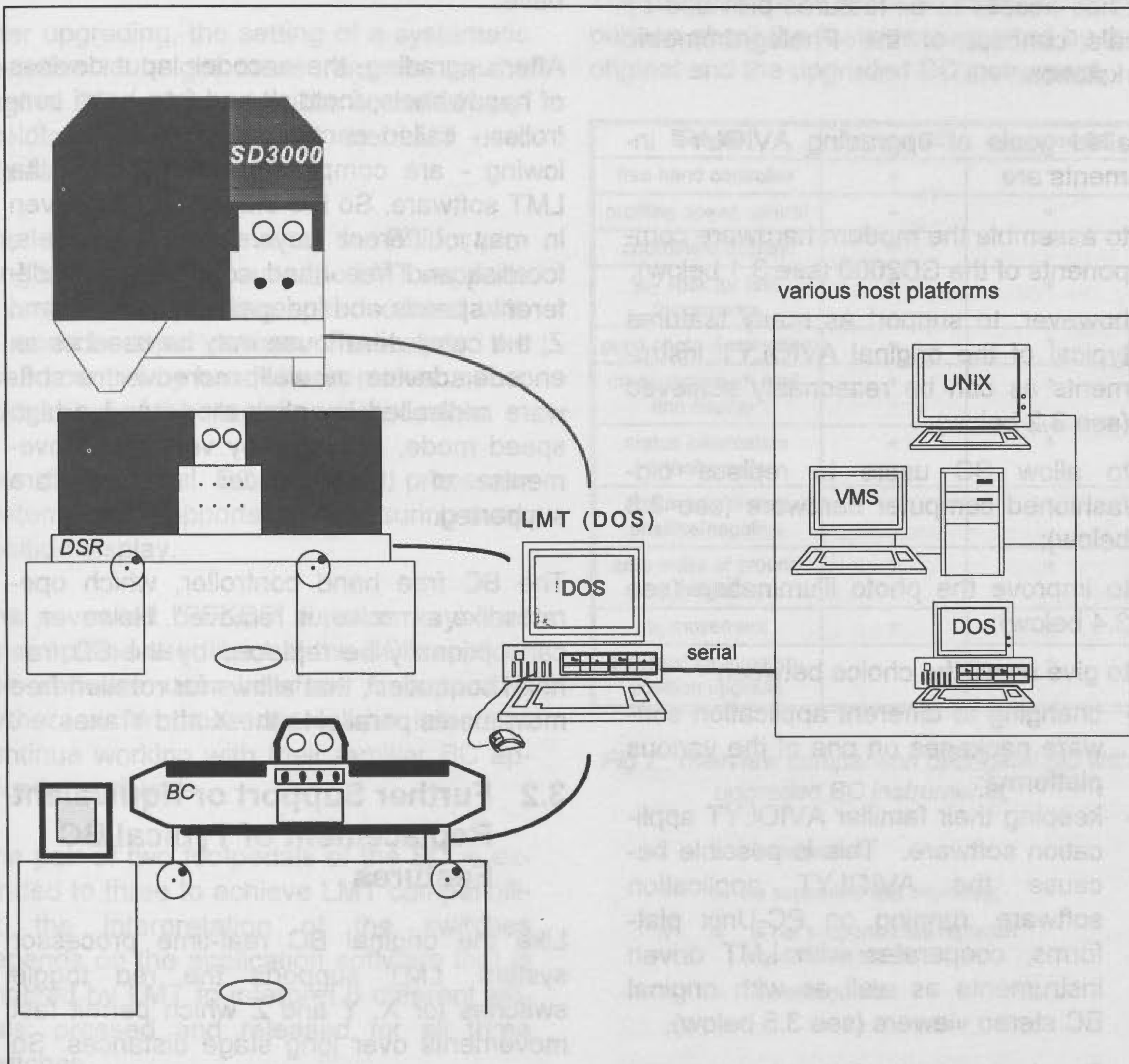


Fig. 1 : Photogrammetric Workstation, supporting 3 instrument lines

3. Details of the AVIOLYT Upgrade Kit

Upgrading an instrument in this context means to make it an LMT driven instrument that has access to all features provided by Leica's concept of the Photogrammetric Workstation.

Detailed goals of upgrading AVIOLYT instruments are

1 to assemble the modern hardware components of the SD2000 (see 3.1 below);

2 however, to support as many features typical of the original AVIOLYT instruments as can be reasonably achieved (see 3.2 below);

3 to allow BC users to replace old-fashioned computer hardware (see 3.3 below);

4 to improve the photo illumination (see 3.4 below);

5 to give users the choice between
- changing to different application software packages on one of the various platforms;
- keeping their familiar AVIOLYT application software. This is possible because the AVIOLYT application software, running on PC-Unix platforms, cooperates with LMT driven instruments as well as with original BC stereo viewers (see 3.5 below).

3.1 Replacement of Hardware Components

The servo control electronics is replaced by just two small electronic cards, one in the BC's cabinet and the other one in a slot of the LMT. The two cards are basically identical to those installed on the SD instruments; these make the stages of the upgraded BC move the same way as the SD stages. The only difference between the new BC cards and the SD cards are some extensions to support BC typical features (see 3.2 below).

The linear encoders for measuring the current stages positions establish the same stage coordinate systems as the SD's have.

After upgrading, the encoder input devices of handwheels, footdisk and free hand controller - called encoder devices in the following - are completely controlled by the LMT software. So the stages can be driven in many different ways: with handwheels, footdisk and free hand controller, with different speeds and independently in XY and Z; the computer mouse may be used as an encoder device as well; moreover, a software controlled joy stick mode and a high speed mode, activated by very fast movements of the encoder devices are supported.

The BC free hand controller, which operates like a mouse, is removed. However, it can optionally be replaced by the SD free hand controller[2], that allows for rotation-free movements parallel to the X and Y axes.

3.2 Further Support or Equivalent Replacement of Typical BC Features

Like the original BC real-time processor system, LMT supports the red toggle switches for X, Y and Z which permit fast movements over long stage distances. So operators can choose between this kind of fast movements and the two provided by LMT anyway (compare 3.1 above).

The coordinate display in front of the stereo viewer is replaced by the LMT's user interface, displaying all components of photo, model or ground coordinates at the same time if required.

The original BC provides a b_y adjustment, performed by an incremental rotary button for slow adjustment and a toggle switch for

[2] This option will not be available for S9-AP upgrades

fast adjustment. Since the LMT functionality enables the operator to control this adjustment by software means (using single stage movements e.g. during relative orientation or data capturing for aerial triangulation), this feature is removed.

After upgrading, the setting of a systematic deviation in height measurements, at the original BC performed by the so called special keys, may be set at the LMT user interface.

The speed control lever of the BC for profiling is removed because the LMT supports a very flexible profiling speed control with all encoder input devices. Particularly, the speed control by handwheels makes a finer adjustment of the stages speed possible.

Like the original BC real-time processor system, LMT supports the measuring mark position display.

The so called "PFKB5" function keyboard for simplified user input at the BC's application software user interface is supported further on for those users who intend to continue working with their familiar BC application software.

The pair of two footpedals of the BC is extended to three to achieve LMT compatibility; the interpretation of the switches depends on the application software that is enabled by LMT to interpret 6 different signals: pressed and released for all three switches.

Like the original BC real-time processor system, LMT sends a message to the host application carrying status information about the BC Ortho/Pseudo lever. Packages running on the host application computer can access this information very easily. However, current versions of application software have to be adapted if this feature is required.

Like the original BC real-time processor system, LMT checks and consideres the status of the optical positive/negative changeover switch, i.e. diapositives and negatives may be restituted in the same way as before. The different positioning of the stages is handled by LMT itself, without affecting the application software at all.

The following table gives an overview comparison about the features supported by the original and the upgraded BC instrument.

feature	original BC	upgraded BC
free hand controller	+	*
profiling speed control	+	+
coordinate display	+	*
joy stick for fast movements	+	+
good photo illumination	+	*
measuring mark position display	+	+
status information ortho/pseudo	+	+
status information positive/negative	+	+
zero index of ground height	+	+
b_y movement	+	o
potential for superimposition upgrade	-	+

Fig.2 : overview comparison of original BC with upgraded BC instruments;

+ = is supported;

* = is supported and improved;

o = is not supported but replaced by different tools;

\- = is not supported

3.3 Computer Replacement

The real-time microprocessor system on the original BC is replaced by an Intel based PC that fits into the cabinet of the BC. If the application host computer is a PC as well, both computers fit into that cabinet. The communication between both platforms is serial.

3.4 Improvement of Photo Illumination

The well proven fluorescent tube illumination type of the BC will not be changed, but the original unit is replaced by an up-to-date unit of the same type, that allows for a 100% disturbance-free adjustment over the whole brightness range.

3.5 Extended Choice of Application Software

As illustrated in Fig. 1, various platforms may be connected to LMT by serial communication. One of these platforms is the Unix-PC, where users of original BC's can run their usual application software, including the PFKB5 support.

However, after upgrading, there is a range of application software provided by Leica that the user could choose from, e.g. with PC-Pro600 the user is able to capture data and do mapping inside the MicroStation[3] CAD system, based on orientations perfomed either with MSU model setup software for aerial photogrammetry, or with MAAS-CR for close range photogrammetry; he could do mapping inside Leica's INFOCAM GIS system, running on DEC VAX-VMS and AXP-VMS.

4. Conclusion

With the new Leica AVIOLYT upgrade kit the user gets new hardware and software for his BC stereo viewer.

The hardware part of the upgrade allows to remove old components that cannot be replaced by identical units any more. This reduces the risk to loose operationality in case a hardware component fails.

The software upgrade integrates the stereo viewer into Leica's concept of the Photogrammetric Workstation, giving the user a bigger choice from various application platforms and software packages, including his familiar AVIOLYT application software.

References:

Cogan, Luis; Hinsken, Ludger 1992: "The Concept of a Photogrammetric Workstation Outlined by the Example of the Leica SD2000". International Archives of Photogrammetry and Remote Sensing VOL XXIX Part B2, Washington 1992

Hinsken, Ludger; Meid, Alfons 1993: "The Use of Leica's Universal Analytical Photogrammetric Workstations for Close Range Applications". In "Optical 3-D Mesurement Techniques II", Wichmann Verlag GmbH, 1993, ISBN 3-87907-254-X

Acknowledgement

I wish to thank Martin Stampfli, electronics engineer at Leica PMU, as well as the Technical Service Group at Leica Heerbrugg, particularly Norbert Lämmer and Cedric Messner, for their excellent contributions to the upgrade development.

3 MicroStation is a trademark of Bentley Systems, Inc. an Intergraph affiliate.

AN EVEN FASTER RANGE SEARCH ALGORITHM FOR MULTI-DIMENSIONAL POINT SETS

Y. C. LEE and BENSON O. AGI

Department of Geodesy and Geomatics Engineering
University of New Brunswick, Fredericton, N. B., Canada E3B 5A3

Phone: (506) 453-5148, Fax: (506) 453-4943, E-mail: yclee@unbmvs1.csd.unb.ca

ABSTRACT

The paper describes a range search algorithm on two-dimensional point sets based on the row-major ordering of points. These algorithms are particularly useful for bathymetric applications where a large number of soundings are often involved.

The algorithm uses a range search strategy developed earlier at the University of New Brunswick but based on the Morton ordering of points. That strategy does not require any spatial indices other than a sorted list of Morton codes. Its performance was found to compare favourably with the fastest range search method on point sets.

It was discovered that the recursive nature of Morton ordering has adverse effects on range search, particularly in terms of multi-dimensional extensions and multi-sided query windows. We will describe these problems and their solutions in this paper. As a result, an even faster range search algorithm with better performance has been developed.

KEY WORDS: Geographic Information Systems, Algorithms, Range Search, Point Sets.

1. INTRODUCTION

The requirement for an increasingly large spatial database is inevitable. While the ability of computers has steadily improved, the number of data sources and densities have also increased [McCormick et al., 1987]. The rate of increase of data volume arising from higher resolutions of data (e.g., new earth resource satellite data) far exceeds the rate of improvement in computer efficiency. Goodchild [1989] has shown that a doubling of spatial resolution produces at least a fourfold increase in data volume, in many cases.

Applications involving large amounts of spatial data cause special concerns in the field of spatial database management systems [Guenther and Buchmann, 1990; Yang, 1992]. The application areas include pure geometric computations, entity relationship representations, resource and administrative management, socio-economic and administrative planning, on local and global scopes.

Operations in a GIS require fast access and retrieval of data in spatial databases. On the other hand, large volumes of data involved in spatial databases cause slow response to user queries for some spatial data. The solution to this problem generates the need to employ optimized data structures and algorithms for spatial indexing and searching [Yang, 1992]. While various search algorithms exist, there is a continual need for improvement as uses of GIS diversify.

Searching in a spatial database is influenced by the type of *ordering* of the data space. The performance of any search algorithm would depend on the optimization which could possibly be applied to that algorithm. *Linear indexing* structures have been favoured for use in GIS because of their advantages [Yang, 1992]: they optimize the storage and query process; they are simple and flexible, allowing for both direct and sequential access to any point; and they allow optimization of algorithms necessary for image database applications [Orenstein and Manola, 1988]. The ordering scheme which has been popular in database applications is the *Morton order*.

A 2-d search algorithm based on the Morton order is compared, in this paper, with a new algorithm based on *row-major ordering*, in terms of its simplicity and performance. Problems identified with the range search algorithm using Morton codes are discussed, and solutions are suggested to overcome these problems. As a result, an extended version of the 2-d range search algorithm to a multi-dimensional solution , using the *layered approach*, has been achieved.

2. SUMMARY OF THE 2-D ALGORITHM BASED ON MORTON ORDER

A detailed description of the range search algorithm based on Morton order has been given by Yang [1992] and Lee and Yang [1993]. The basic concept on which that algorithm was developed is that: *over-searches* are reduced by keeping the search within the query window. The moment a data point is found to be outside the window, the next entry point of the window is accessed by searching along the edges of the window. A data point immediately after that is called the Forward Point (FP). Some optimizations, based on the properties of Morton curve, are then applied, to further reduce the search over-head. The major advantage of this strategy is that it does not need additional data structures to support the search, and the search time depends on the query window with less influence from the size of the database.

The performance of that algorithm was judged to largely depend on how fast FP could be found. The preprocessing required to create a sorted list of Morton codes is of O(NLogN), where N is the number of data points. The storage complexity is O(N). The processing time required to find an FP, and retrieve the data at or after FP, is O(KLogL + KLogN); where L is the average length of an edge in resolution units, and K is the number of times the Morton curve exits the window. On a fully-populated space, N > L, and the search over-head is bounded by KLogN. K depends on the size and location of the query window. The time it takes to retrieve I data points within a window of L by L units is O(I + KLogL + KLogN).

3. PROBLEMS WITH RANGE SEARCH BASED ON MORTON ORDER

The following problems have been identified with the search algorithm based on Morton order:

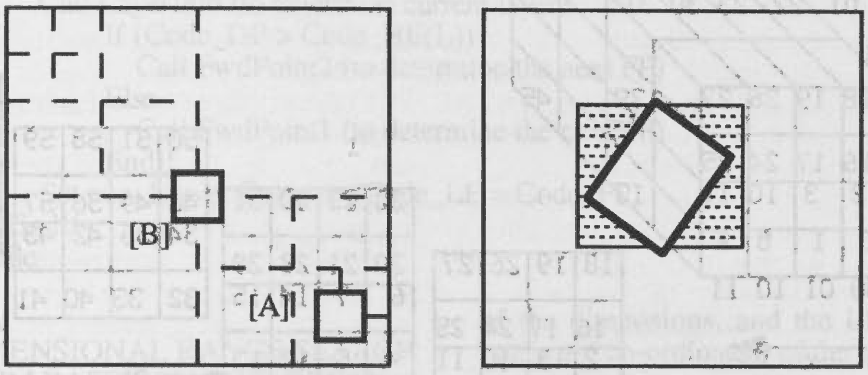
1. The time for finding a forward point (FP). This depends on the size and location of the query window within the application space. The recursion that generates the Morton codes is quadrant-based. Therefore, as the query window cuts across different quadrants, the processing of the forward points take on different characteristics. For example, for a given 2-d space as in Figure 1a, the query time for the same size of a window in location [A] is expected to be significantly lower than that in location [B] which cuts across higher levels of quadrants.

2. Costly overhead for rotated or generally irregular window. The 2-d search algorithm works only for a query window which sides are orthogonal to the co-ordinate axes. This is because its efficiency depends on the properties of Morton order, namely: (a) the Morton codes on lines parallel to the axes of the space are sorted; and (b) for any rectangular window orthogonal to the axes, the south-west corner (SW) has the smallest Morton code, and the north-east corner (NE) has the greatest code. This constraint necessitates the forming of bounding rectangles over query windows which have irregular sides, or rectangular windows which are not orthogonal to the co-ordinate axes.

As a result of this constraint, over-searches corresponding to the shaded area of Figure 1b arise. The accuracy of the search could be jeopardized, and any attempt to improve the accuracy may be at the expense of search time.

3. Extension of the concept to multi-dimensional range search. It seems possible to extend the 2-d algorithm to involve domains of n-d spaces. In this case, the file is a collection of records, each of which is identified by an ordered n-tuple of keys $\{W_1, W_2, ..., W_n\}$ which can be viewed as a point in an n-d Cartesian space. The Cartesian space formulation is the abstraction of a variety of important applications, often referred to as *multikey search* [Knuth, 1973; Yang, 1992]. In the Cartesian space, aspatial problems can bear geometric significance if the record attributes are regarded as co-ordinates and the n-values for each record as representing a point in an n-d co-ordinate system.

One can draw an analogy between the 2-d and the n-d case. An n-d query window is a solid figure composed of hypersurfaces. The analogy is to order the n-d space and search in the hypersurfaces (instead of along the edges, in the 2-d case) for the next data point in the window. The optimization of this search would involve, in the least, the determination of all possible faces in which the entry point could be found, and searching in these faces to find the first points. Doing this would require complex analysis, because of the complexity of hypersurfaces. It has already been shown in the 2-d algorithm that the cost of the search is highest in the optimization which finds the first points. With the complexity of hypersurfaces, the time for finding the FP's may be extremely high.



(a) Influence of Location of Query Window

(b) Over-head in a Rotated Window

Fig. 1

These problems, that have been identified with range search based on Morton ordering, generate a need to explore some other ordering schemes which has, at least, the simplicity of Morton order, and at the same time can offer some advantages for extending the 2-d search concept to n-d search.

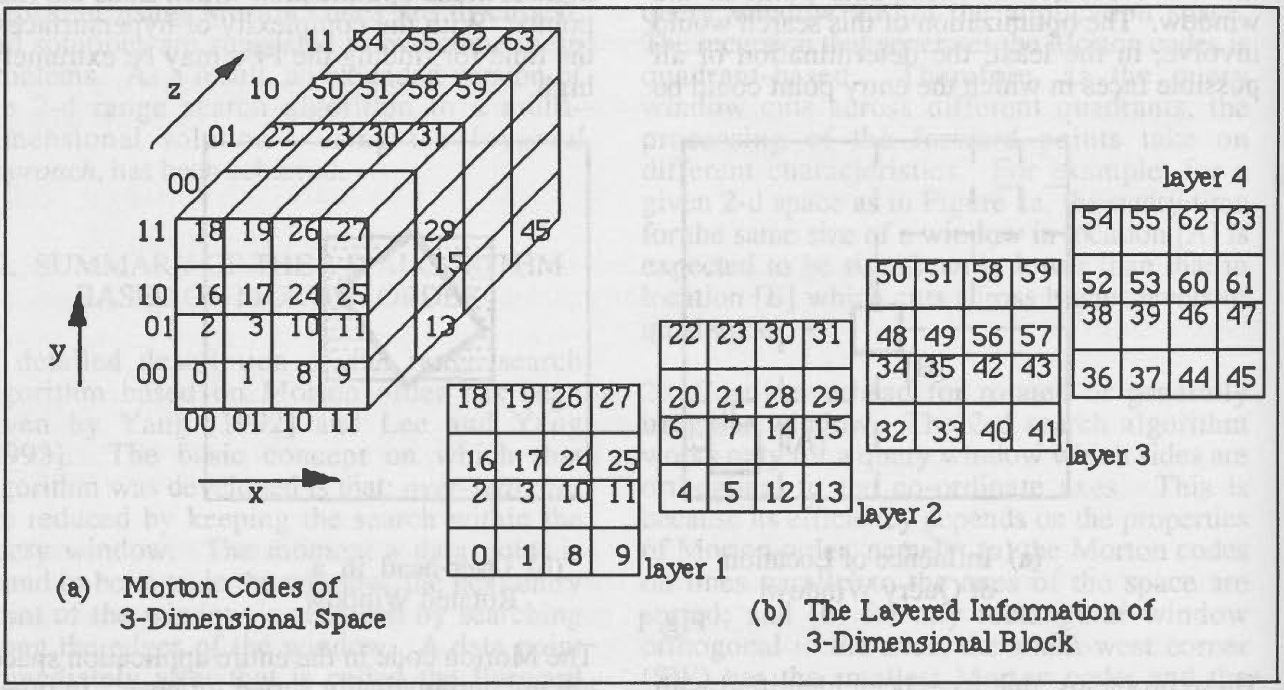
4. THE LAYERED APPROACH SOLUTION AND SPACE ORDERING

To benefit from the simplicity of the 2-d range search algorithm, the concept can be extended to the n-d case. The n-d space is treated as a collection of 2-d layers. Points in each layer are consecutively coded, starting from the south-west corner of the basic or first layer (the x-y plane). All subsequent layers are then a projection of each layer on the x-y plane. The coding of points within these layers is done in such a way that the south-west corner of the first layer is given the code 0, and all the points within the layer are completely ordered until the north-east corner point, which has the maximum code in that layer. The code of the south-west corner of any succeeding layer is equal to the maximum code in the previous layer plus 1.

The Morton code in the entire application space are not automatically sorted by layer. In other words, a Morton curve does not completely traverse a layer before visiting another one. Instead it goes in and out of any given layer, depending on the size of the resolution units in every dimension greater than 2. This fact is demonstrated in Figure 2, using a 3-d space as an example. As a consequence, it is difficult, if at all possible, to apply the same technique of search optimization on Morton codes.

The *row order* which does not seem to impose the same restriction in coding and optimization has been investigated, as an alternative to Morton order, in the solution of n-d range search. The row order completely traverses a layer before going into another layer (see Figure 3). Other properties of the row order are that: the codes are always sorted in increasing y-co-ordinates, even on profiles not parallel to the co-ordinate axes; it is also always sorted in x-co-ordinates, given any constant y-value.

Due to the properties of row order, it seems to allow better optimization of n-d range search. Apart from being compatible with the layered approach, it also makes the determination of the first points (FP's) much easier than in the case of Morton order. (Recall that the most time is spent on the determination of the FP's in the 2-d algorithm using Morton order.) This is because, the problem of crossing bridges (as in Morton order) is completely absent in row order. Also, the analysis required in determining which possible edges of the query window to search for the entry point is not necessary for orthogonal windows, since the entry point is always on the west edge of the window. With the row order, the search time does not seem to depend on the location of the window within the application space. It only depends on the magnitude of the window.



(a) Morton Codes of 3-Dimensional Space

(b) The Layered Information of 3-Dimensional Block

Figure 2: Morton Ordering of a 3-Dimensional Block

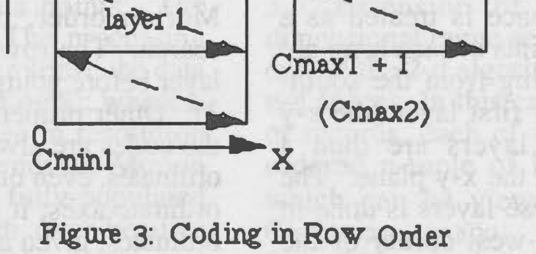


Figure 3: Coding in Row Order

The properties of row order also seem to enable better opitimization for rotated and irregular windows. Since the row codes are always sorted in increasing y, for any given layer, and the layers are also sorted in increasing dimensions, a search algorithm using row order would seem to be less complex and more efficient than that using Morton order. For one thing, there would be no need for forming bounding rectangles, thereby eliminating the oversearches and possible inaccuracies which would otherwise be the case.

LISTING 1: Pseudocode of the n-d search algorithm for a regular window

```
Begin
 Compute the codes of the QW extremes {Code_LE, Code_HE}
   Find the next data point, DP (Code_DP), at or after Code_LE
      While { (Code_DP < Code_HE) and end of data point list not reached}
         Call IncludeTest
            If (Code_DP is IN) Then Report DP
               Take next DP sequentially
            Else (Code_DP is OUT)
               Call LayerInfo (to determine current layer)
                  If (Code_DP > Code_HE(L))
                     Call FwdPoint2 (to determine the next FP)
                  Else
                     Call FwdPoint1 (to determine the next FP)
                  EndIf
               Set new Lower Extreme (Code_LE = Code_FP)
            EndIf
      EndWhile
EndBegin.
```

5. AN N-DIMENSIONAL RANGE SEARCH ALGORITHM BASED ON ROW ORDER

The pseudocode of an n-d range search algorithm for an orthogonal query window is given in LISTING 1. The following explanations are necessary for understanding the pseudocode:

QW is the query window.

LE, HE are the Lower Extreme (south-west corner point) and Higher Extreme (north-east corner point), respectively, of the query window.

DP, FP stand for Data Point and Forward Point respectively.

HE(L), NE(L) are the Higher Extreme and the North-East point (the highest point) of any arbitrary layer, L.

Code_p is the code of any given point, p.

The function, IncludeTest tests if a given point is inside the window, while the function LayerInfo determines the next layer with its relevant homolog points (those corresponding geometrically to the LE, HE, and NE of any given layer L), whenever a data point is found to be outside the window (this information is relevant for the determination of the next forward point). The function FwdPoint1 determines the next FP when the search is still within the same layer, while FwdPoint2 determines FP when a data point which has been found to be out of the window is also found to be in a different layer.

Given are the sorted points with codes precomputed, the highest resolution unit in each of the dimensions, and the LE and HE (each being n-d co-ordinates) of the query window.

5.1 Performance of the new algorithm in 2-d

The preprocessing required to create a sorted list of row codes takes O(NLogN) time, where N is the number of data points. The storage complexity is O(N). Each time the search exits the window, it requires a constant time to compute the next FP, and a time of O(LogN) to retrieve FP or the next data code from the list. Hence, the total determination and retrieval of FP's is of O(Yrange + YrangeLogN) where Yrange is the difference in resolution units between the y-values of LE and HE of the query window. This is the worst case for a fully-populated space. The total search time for retrieving I data points in a given window is therefore O(I + Yrange + YrangeLogN).

6. A COMPARISON OF THE 2-D MORTON ORDER AND ROW ORDER SEARCH ALGORITHMS

6.1 Theoretical comparison

The performance of the Morton order search is dependent on both the size and location of the query window, whereas, the row order search is only dependent on the size of the window. Also, the determination of FP's in row order does not seem to be as critical as that of Morton order search.

The time complexities in preprocessing and storage are the same in both orders, but there is a significant difference in the processing times. While the processing time of Morton order search is O(I + KLogL + KLogN), that of row order is O(I + Yrange + YrangeLogN). Then, what determines which algorithm is better very much depends on the evaluation of the terms KLogL versus Yrange, and the constants K versus Yrange. Since there does not seem to be any objective way to evaluate these parameters theoretically, an experiment has been conducted to compare their practical performances. What is significant to note is that the determination of FP's in the row order search is consistently dependent on the y-range of the query window.

6.2 Experimental comparison

The new range search algorithm has been implemented in C, using a fully-populated 2-d space, defined over 1024 by 1024 points. It has been run on the same platform, a SUN SparcStation SLC. as used in the previous test using the Morton order.

Statistics are generated for CPU times for both algorithms, using three varying configurations of query windows, namely:

(a) C-configuration: a query window of size 300 by 300 resolution units, moved over nine different locations of the data set (see Fig. 4a). The statistics comparing both methods are given in Table 1.

(b) Z-configuration: the size of the query window is varied by reducing it to 100 by 100 units while maintaining the same locations as before (see Fig. 4b). The statistics are given in Table 2.

(c) H-configuration: the shape of the query window is varied and tested at different locations (Fig. 4c). The statistics are in Table 3.

6.1 Analysis of experimental results

1. It can be concluded from columns (6) and (7) of each of the three tables that the algorithm based on row order performs much better than that based on Morton order. Generally, the CPU time for the search in Morton order takes about 2.5 times the time for the search based on row order, for the same number of window points (i.e., same size of query window), and at the same location within the application space.

2. As theoretically expected, the result of the experiment shows that a search on Morton codes performs worst at the centre of the application space, because of the quadrantic nature of Morton curve. This is evident in column (6) at each of the central locations C5 and Z5. The time for C8 looks like an exception, and may be partly because its location also cuts across two major quadrants, and also that higher digits are involved in the handling of co-ordinates and codes of points.

3. Column (7) of both Tables 1 and 2 shows that the search time using row ordering does not depend on the location of the query window. The time, however, generally increases with increasing y. This can be partly due to the y-range factor in the theoretical complexity, and also partly to the fact that higher digits are being manipulated with increasing y-coordinates.

4. It is observed from Tables 1 and 2 that reducing the size of the query window by nine times also reduces the query time by similar amount, in both Morton and row orders. Hence, the factor by which row order out-performs Morton order is still maintained.

5. It is observed from Table 3 that variation of the shape of the query window maintains the same relationship in performance between the search over Morton and row codes.

7. CONCLUSION

An n-d range search algorithm, based on row order, has been developed for the special case of orthogonal query windows. It has been implemented using test data in 2-d space. A theoretical analysis has been carried out to show that row ordering may be more suitable than Morton ordering of space, in the extension of an existing 2-d range search algorithm to n-d, because of some constraining properties of the Morton curve. An experimental comparison of 2-d range search over both ordering has shown that row order out-performs Morton order by about 2.5 times in query response time.

The extension of the new range search algorithm to rotated and irregular query windows, with the full implementation of the n-d version, is in progress.

8. REFERENCES

Goodchild, M.F. (1989). "Tiling of Large Geographical Databases." *Proceeding of "Design and Implementation of Large Spatial Databases". First Symposium SSD'89*, Santa Babara, U.S.A., pp. 138-146.

Guenther, O. and A. Buchmann (1990). "Research Issues in spatial Databases." *SIGMOD Record*, vol.19, no.4, December, 61-68.

Knuth, D.E. (1973). "The Art of Computer Programming. Volume III/ Sorting and Searching." Addison-Wesley, Reading, Mas.

Lee, Y.C. and W.P. Yang (1993). "A Fast and Simple Range Search Algorithm for Point Sets." *Proceedings of the Canadian Conference on GIS - 1993*, 975-987.

McCormick, B.H.; T.A. DeFanti and M.D. Brown (1987). "Visualization in Scientific Computing." *Computer Graphics*, vol.21, no.6, ACM SIGMOD: New York.

Orenstein, J.A. and F.A. Manola (1988). "PROBE Spatial Data Modeling and Query Processing in an Image Database Application." *IEEE Transaction on Software Engineering*, vol.14, no.5, pp. 611-629.

Yang, W.P. (1992). "A New Range Searching Algorithm for Large Point Databases." Unpublished Masters Thesis, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada.

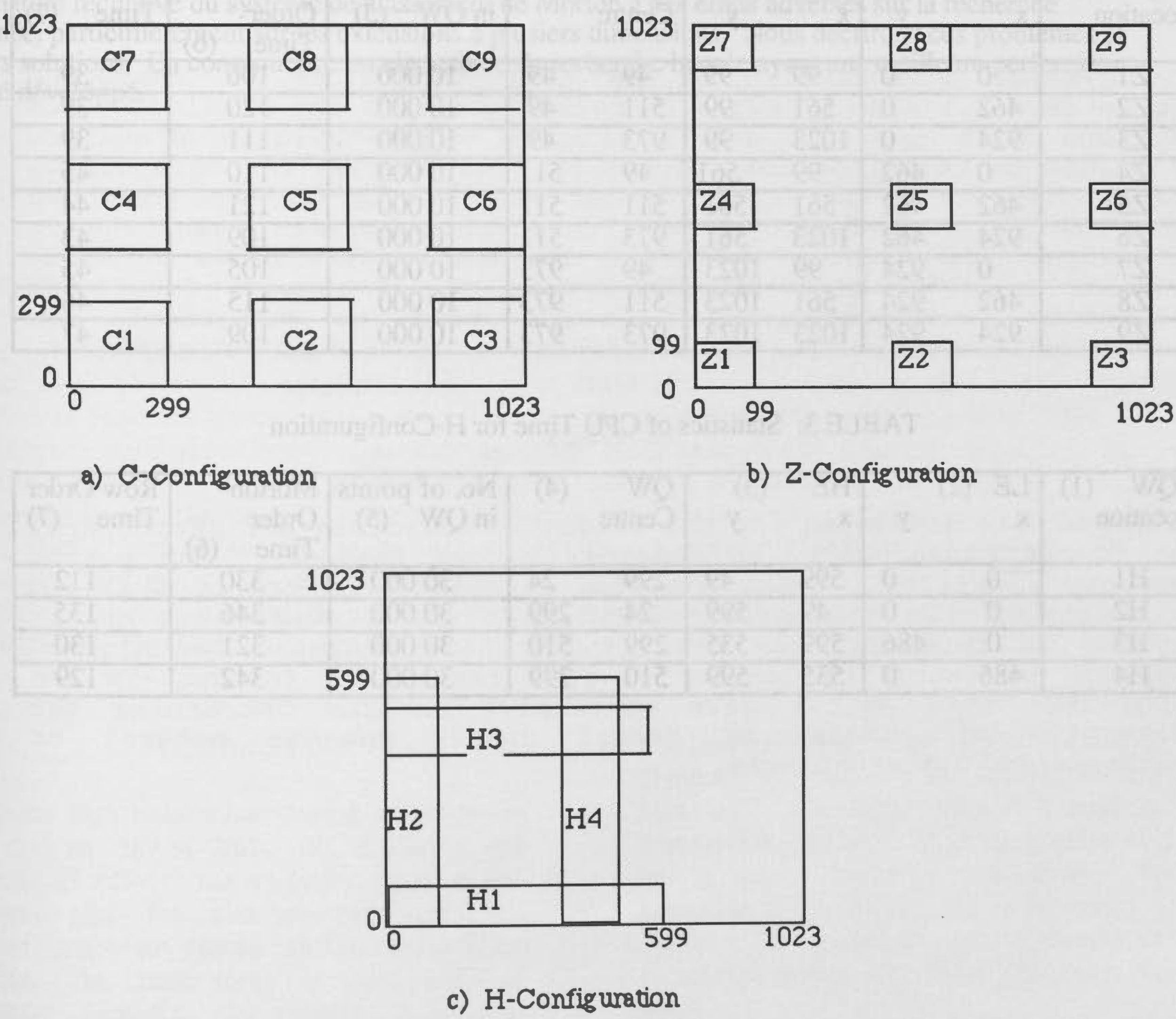


Fig. 4: Variation of Query Window Configuration

TABLE 1: Statistics of CPU Time in Clock Cycles for C-Configuration

QW (1) location	LE (2) x	y	HE (3) x	y	QW (4) Centre		No. of points in QW (5)	Morton Order Time (6)	Row Order Time (7)
C1	0	0	299	299	149	149	90 000	924	369
C2	362	0	661	299	511	149	90 000	981	364
C3	724	0	1023	299	873	149	90 000	952	371
C4	0	362	299	661	149	511	90 000	945	389
C5	362	362	661	661	511	511	90 000	986	392
C6	724	362	1023	661	873	511	90 000	957	397
C7	0	724	299	1023	149	873	90 000	941	398
C8	362	724	661	1023	511	873	90 000	992	401
C9	724	724	1023	1023	873	873	90 000	951	400

TABLE 2: Statistics of CPU Time for Z-Configuration

QW (1) location	LE (2) x	y	HE (3) x	y	QW (4) Centre		No. of points in QW (5)	Morton Order Time (6)	Row Order Time (7)
Z1	0	0	99	99	49	49	10 000	106	39
Z2	462	0	561	99	511	49	10 000	120	39
Z3	924	0	1023	99	973	49	10 000	111	39
Z4	0	462	99	561	49	511	10 000	110	45
Z5	462	462	561	561	511	511	10 000	121	44
Z6	924	462	1023	561	973	511	10 000	109	43
Z7	0	924	99	1023	49	973	10 000	105	45
Z8	462	924	561	1023	511	973	10 000	115	46
Z9	924	924	1023	1023	973	973	10 000	109	47

TABLE 3: Statistics of CPU Time for H-Configuration

QW (1) location	LE (2) x	y	HE (3) x	y	QW (4) Centre		No. of points in QW (5)	Morton Order Time (6)	Row Order Time (7)
H1	0	0	599	49	299	24	30 000	330	112
H2	0	0	49	599	24	299	30 000	346	135
H3	0	486	599	535	299	510	30 000	321	130
H4	486	0	535	599	510	299	30 000	342	129

UN ALGORITHME POUR LA RECHERCHE BORNÉE (RANGE SEARCH) PLUS RAPIDE POUR DES ENSEMBLES DE POINTS À PLUSIEURS DIMENSIONS

RÉSUMÉ

Cet article décrit un algorithme pour la recherche bornée (range search) sur un ensemble de points en 2 dimensions basé sur le principe de rangée-majeure. Ces algorithmes sont très utiles pour des applications en bathymétrie dû au grand nombre de données.

L'algorithme utilise la stratégie de "recherche bornée" qui a été dévelopée à l'Université du Nouveau-Brunswick et basée sur la méthode de classement des points de Morton. Il n'y a aucun besoin d'indice mais seulement une liste classée de codes de Morton. Sa performance est comparable aux méthodes de recherche bornées les plus rapides.

La nature récursive du système de classement de Morton a des effets adverses sur la recherche bornée, particulièrement sur les extensions à plusieurs dimensions. Nous décrirons ces problèmes et leurs solutions. En conséquence un algorithme de recherche bornée ayant une meilleure performance a été développé.

INVENTAIRE ET CARTOGRAPHIQUE AUTOMATIQUES DE LA RESSOURCE FORESTIÈRE À L'AIDE DES IMAGES DE TÉLÉDÉTECTION

François Cavayas et Stéphane Chalifoux

Laboratoire de télédétection et de cartographie numérique
Département de géographie,
Université de Montréal
C.P. 6128, Succ. "A"
Montréal (Québec)
H3C 3J7
Tél. : ; Téléc. :

RÉSUMÉ

Un système d'inventaire et de cartographie de la ressource forestière à l'aide d'images de télédétection a été développé. Ce système fait appel à des techniques algorithmiques d'analyse d'images de télédétection guidées par les données cartographiques existantes sur le milieu forestier étudié. Un système expert à base de règles est chargé de l'interprétation des résultats de l'analyse numérique. Appliqué à des images satellites de haute résolution (TM-Landsat et HRV-SPOT), ce système a donné des résultats concluants quant à la détection des changements des peuplements forestiers dûs à des facteurs naturels ou anthropiques.

ABSTRACT

A system was developed for the automatic inventory and cartography of forest resources using remotely sensed images. The system employs algorithmic image analysis techniques guided by existing map data on the studied forestry area, and a rule based expert system to interpret the results of these analysis techniques. Applied to satellite imagery of fine resolution (Landsat-TM and SPOT-HRV) this system gave conclusive results concerning the detection of changes of forest stands due to natural or anthropogenic disturbances.

PERFORMANCE PREDICTION OF AVNIR BY A SIMULATOR

HAJIME KOSHIISHI

Remote Sensing Technology Center
7-15-17, Roppongi, Minatoku, Tokyo, Japan

MASAO NAKA

National Aerospace Laboratoly

YOSHIYUKI KAWATA

Kanazawa Technology Institute

YOSHITAKA IWATA

Fujitu F.I.P Co.Ltd

ISPRS Commission II, Working Group 2

KEY WORDS: Imaging spectrometer, Pushbroom scan, Simulation, ADEOS, AVNIR

ABSTRACT

The AVNIR(Advanced Visible Near Infrared Radiometer) is one of core sensors of japanese next earth observation satellite ADEOS. The AVNIR has linear array for each multispectral band and panchromatic band, and uses the pushbroom scheme for scanning. The pushbroom scan has a sampling process in cross track scan as the inherent property. In along track scan, there is no such a inherent sampling process, but there will be a image degradation by the satellite motion. In such a high resolution spectrometer, the atmospher has significant effect on the spatial resolution. A digital simulator for performance analysis and design of imaging spectrometer was developed. The performances of AVNIR affected by such a sampling process, degradations by satellite motion and atmosphere were studied by the simulator and the results are presented.

1.INTRODUCTION

Recent high resolution imaging spectrometer , such as HRV of SPOT, OPS of JERS-1, and MESSR of MOS-1, has a linear array on the focal plane for each spectral band, and uses pushbroom scheme in the cross track scan. The linear array includes a lot of sensor elements, for example, 6,000 elements for HRV and 4,000 elements for OPS. The linear array is put on the focal plane of optics in cross track direction, and the electron charge of each sensor element corresponds to radiance is scanned electronically. Therefore, here is a sampling process inherently. Aliasing ploblem will be in such a pushbroom scan system. For reduction the aliasing, it is necessary to increase the number of sensor element and to cut the higher excessive frequency responce.

In the along track scan, there is no such a alias problem, but the imaging data is degraded by the motion of platform.

The atmosphere has some lens effect and has own MTF characteristic. The image of object on the ground is degraded by this atmospheric effect and the spatial resolution will be decreased in higher frequency domain, and the contrast of image decreases.
Another important problem is a decrease of spatial resolution by a cross talk between the sensor elements. This cross talk is large in long wave length, and spatial resolution power is lower in longer wave length.
To do analyze these problems and to evaluate the performance of imaging spectrometer, it is necessary to simulate the imaging operation of the spectrometer in detail.

2.OPERATION MODEL OF IMAGING SPECTROMETER

To analyze the operational behaviors of imaging spectrometer in detail, the operation model was originated at first. In followings, both of the pushbroom scan and the whiskbroom scan are considered.
Here, following symbols are defined. IS is abbreviation of imaging spectrometer.

$h(x,y)$: point spread function(PSF) of IS
$r(x,y)$: radiance distribution of ground
$s(x,y)$: sampling function of the IS
$g(x,y)$: sampling output of the IS
$g_p(x,y)$: hold output of pushbroom IS
$g_w(x,y)$: hold output of whiskbroom IS
x,y : spatial coordinate
p,q : pitch of sensor element
$Box(p,l)$: hold function of width p and height l.
l : along track flight distance for integration time
l_o: initial point of the integration

$H(u,v)$: Fourier transform of $h(x,y)$
$R(u,v)$: Fourier transform of $r(x,y)$
$S(u,v)$: Fourier transform of $s(x,y)$
$G(u,v)$: Fourier transform of $g(x,y)$
$G_p(u,v)$: Fourier transform of $g_p(x,y)$
$G_w(u,v)$: Fourier transform of $g_w(x,y)$
u,v : spatial frequency
u_s : $=1/p$, sampling picth
v_s : $=1/q$, sampling picth
$p\cdot\sin\pi up/\pi up$: Fourier transform of hold function $Box(p,l)$
$l\cdot\sin\pi vl/\pi vl$: Fourier transform of hold function $Box(l,l)$

2.1 Operation model in cross track

In the case of pushbroom scan, the output of linear array is a pulse train and expressed as follows.

$$g(x)=\{r(x)*h(x)\}\cdot\sum_{n=-\infty}^{\infty}\delta(x-np) \quad (1)$$

And the hold output of pushbroom IS is,

$$g_p(x)=\left[\{r(x)*h(x)\}\cdot\sum_{n=-\infty}^{\infty}\delta(x-np)\right]*Box(p,l) \quad (2)$$

In the frequency domain, these expressions are as follows.

$$G(u)=\{R(u)\cdot|H(u)|\}*u_s\cdot\sum_{n=-\infty}^{\infty}\delta(u-nu_s) \quad (3)$$

And the Fourier transform of hold output is

$$G_p(u)=\left[\{R(u)\cdot|H(u)|\}*\sum_{n=-\infty}^{\infty}\delta(u-nu_s)\right]\cdot\sin\pi up/\pi up \quad (4)$$

In the case of whiskbroom scan, one sensor element scans continuously. Then the output is,

$$g(x)=\{r(x)*h(x)\}\cdot\int_{-\infty}^{\infty}\delta(x-p)dp$$
$$=r(x)*h(x) \quad (5)$$

In the frequency domain,

$$G_w(u)=R(u)\cdot|H(u)| \quad (6)$$

2.2 Operation model in along track

In the case of pushbroom scan, the signal is integrated for a while and the spatial resolution is degraded. The output $g(y)$ is given by,

$$g_p(y)=[r(y)*h(y)]*Box(l,l_o,l) \quad (7)$$

In frequency domain, $g(y)$ is transformed to

$$G_p(v)= [R(v) \cdot | H(v) |] \cdot \qquad$$
$$| \cdot \sin \pi vl / \pi vl \qquad (8)$$

In the case of whiskbroom scan, several sensor elements are arranged in along track . Then the sampling process is included. The output is given as follows.

$$g_w(y)= [r(y) * h(y)] \cdot \sum_{n=-\infty}^{\infty} \delta(y - np) \qquad (9)$$

In the frequency domain,

$$G_w(v)= [R(v) \cdot | H(v) |] * \qquad$$
$$v_s \cdot \sum_{n=-\infty}^{\infty} \delta(v - nv_s) \qquad (10)$$

3. THE SIMULATOR

This simulator is constructed on a work station in National Aerospace Laboratory. The simulator is consisted of computer subsystem and dispray subsystem. Features of this simulator are,

(1) flexibility of display subsystem
(2) changeability of simulation contents
(3) simplicity of function addition

The simulation is conducted as a train of subroutines. Table 1 shows the subroutines.

Table 1. Subroutines of the simulator

Title	Function
gscene	generates 19 input scenes
gpsf	generates 15 PSF
wvdisp	graphic dispray of data
power	computation of power
cmfpsf	computation of MTF
convlt	computation of convolution
and other 5 subroutines	
move etc.	data transfer between areas
add etc.	arithmetical operation
sin etc.	trigonometrical function
exp	exponential operation
sqrt	computation of square root
and other 25 operations	

The simulation of imaging operation of IS is conducted as following block diagrams. Fig.1 shows the case without atmospheric effect, and Fig.2 shouws the case with atmospheric effect.

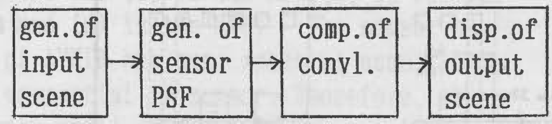


Fig.1 Simulation of the case without atmospheric effect

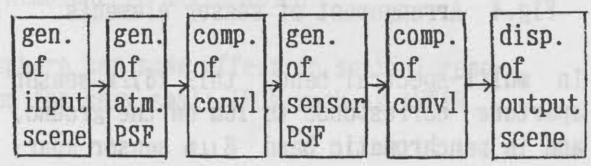


Fig.2 Simulation of the case with atmospheric effect

4. OUTLINE OF AVNIR

The outlook ofAVNIR is shown in Fig.3 The collecting optics is a sort of Schmidt type

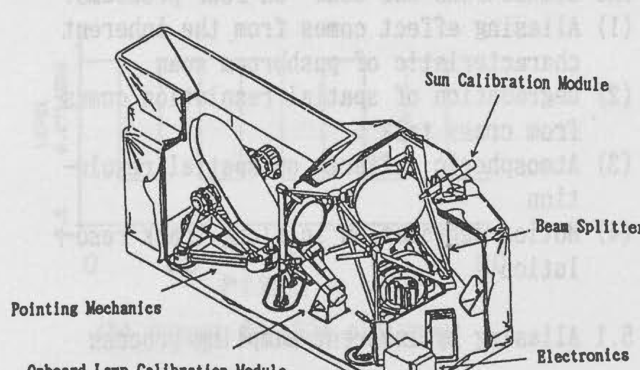


Fig.3 Outlook of AVNIR

catadi-optoric system. A linear array with 5,000 sensor elements is used for each multi-spectral band, and a linear array with 10,000 sensor elements is used for the panchromatic band. The dimension of a sensor element for multi-spectral band is 16μm x 16μm, and the dimension of a sensor element for panchromatic band is 8μm x 8μ

m. The arrangement of the these elements on the focal plane is shown in Fig.4

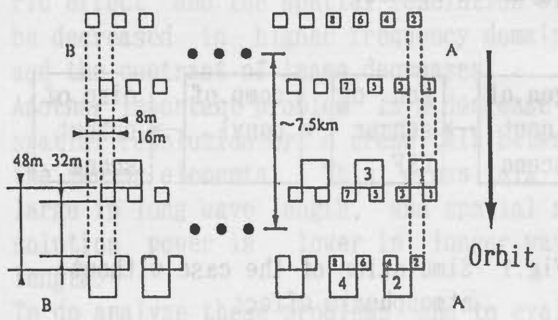


Fig.4 Arrangement of sensor elements

In multi-spectral band, this 16μm sensor aparture corresponds to 16m on the ground, and in panchromatic band 8μm sensor aparture corresponds to 8m on the ground.

5. SIMULATION

The simulation was conducted with a posturated point spread function(PSF) of AVNIR : that is the PSF is assumed to be a box type . Then the moduration transfer function(MTF) of the sensor element is to be a sinc function.

The simulations was done on four problems:

(1) Aliasing effect comes from the inherent characteristic of pushbroom scan
(2) Degradation of spatial resolution comes from cross talk
(3) Atmospheric effect on spatial resolution
(4) Motion degradation in along track resolution

5.1 Aliasing by inherent sampling process

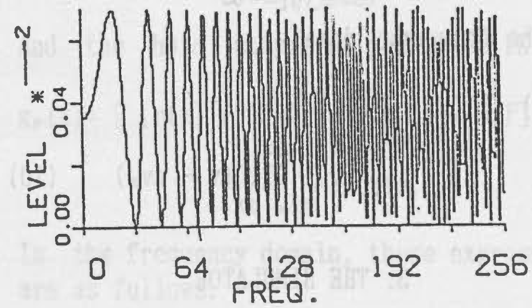
In this simulation, linear zone plate(LZP) was used as the input. This is a modulation of circular zone plate used for TV resolution test. Radiance level of LZP r(x) changes with coordinate x as follows.

$$r(x) = \cos (x^2) \quad (11)$$

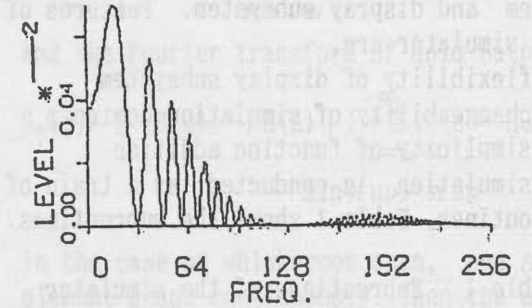
And the spatial frequency U(x) is expressed in next formula.

$$u(x)= [\{x^2 + 2\pi\}^{1/2} + x] / 2\pi \quad (12)$$

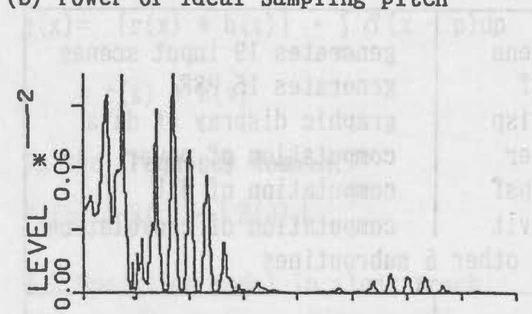
An example is shown in Fig.5 In this Fig. (a) is the power of LZP input. (b) is the out-put power of ideal case in which pitch of sensor element is very short, sampling frequency is very high and no aliasing. (c) is output power of usual case in which pitch of sensor element is equal to the sensor dimension. There is some bump of power in (c) because of aliasing.



(a) Power of LZP input



(b) Power of ideal sampling pitch



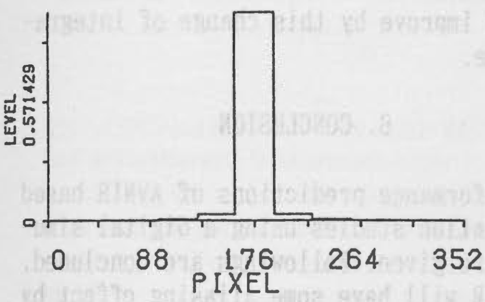
(c) Power of usual sampling pitch

Fig.5 Aliasing effect of sampling pitch

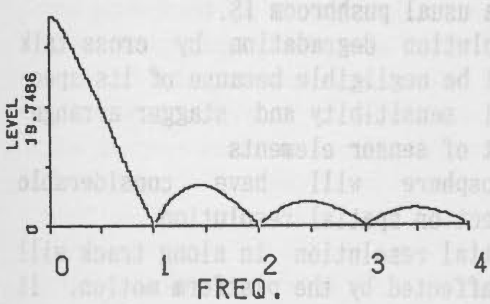
Although this simulation was conducted with a posturated PSF, the essence is same and there will be some amount of aliasing in output.

5.2 Degradation of resolution by cross talk

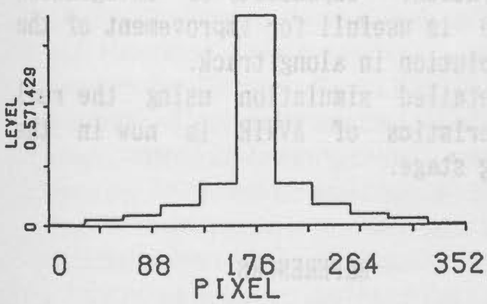
Usualy a linear array has cross talk. The cross talk increases with wave length.



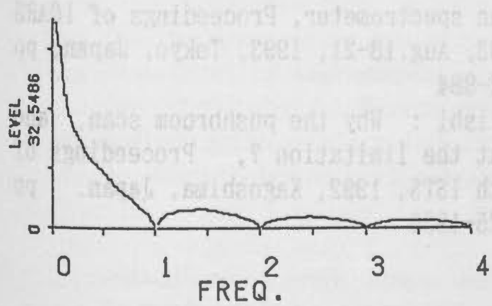
(a) PSF for 550nm



(b) MTF for 550nm



(c) PSF for 850nm



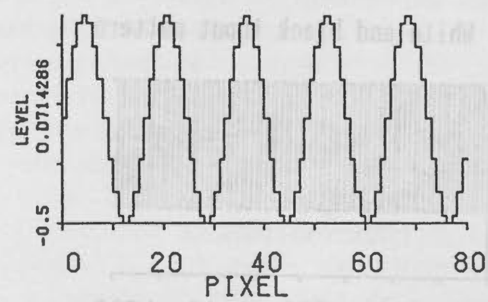
(d) MTF for 850nm

Fig.6 Degradation of MTF by cross talk

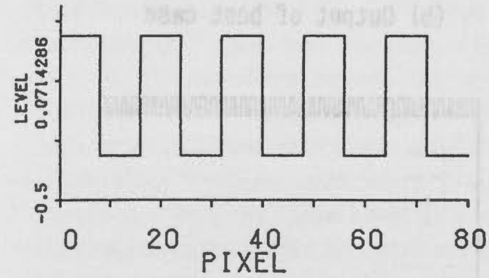
The spatial resolution will be degraded by the cross talk. Fig.6 shows the degradation of MTF with the cross talk. This is not an AVNIR data, but a typical one. (a) is PSF for 550nm and (b) is the MTF.(c) is PSF for 850nm and (d) is the MTF. The sensor element of AVNIR has same spectral sensitivity with commertial TV sensor. Therefore, there is very small cross talk in near infrared wavelength. Besides, stagger arrangement of sensor elements as Fig.3 has good effect on cross talk.

5.3 Atmospheric degradation

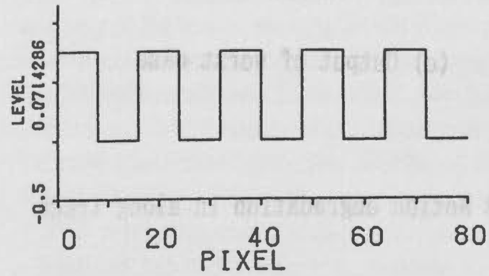
Atmosphere has some effect on spatial resolution for very small IFOV such as AVNIR.



(a) Sin wave input pattern



(b) Output for 50m sin wave

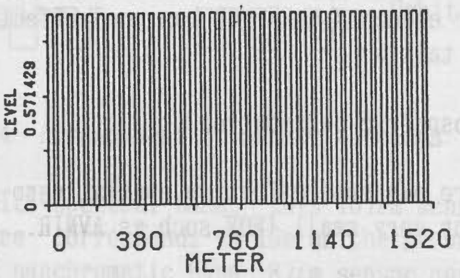


(c) Output for 16m sin wave

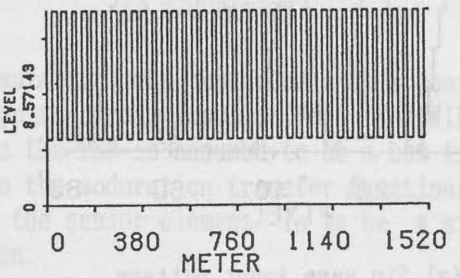
Fig.7 Atmospheric effect on resolution

Fig.7 shows the degradation of spatial resolution by atmosphere. (a) is a black and white sine input scene. (b) is the output for 50m black and white input and (c) is the output for 16m black and white input scene. The atmosphere has serious effect on high resolution imaging.

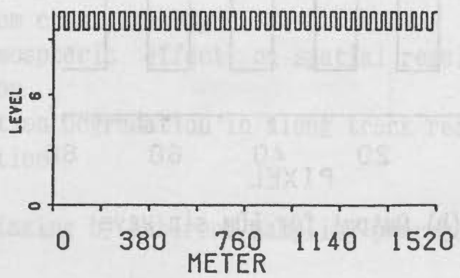
5.4 Motion degradation in along track



(a) White and black input pattern



(b) Output of best case



(c) Output of worst case

Fig.8 Motion degradation in along track

In along track imaging, the signal from the ground is integrated during flight. Therefore, spatial resolution will be degraded by this motion. The amount of the degradation depends on start point of the integration. Fig.8 shows the motion degradation. (a) is the white and black input scene. (b) is the best case and (c) is the worst case. The difference of best and worst is determined by the start point of integration. AVNIR can change the integration time. Resolution in along track be able to improve by this change of integration time.

6. CONCLUSION

Some performance predictions of AVNIR based on simulation studies using a digital simulatior are given. Followings are concluded.

(1) AVNIR will have some aliasing effect by the inherent sampling process as like as a usual pushbroom IS.
(2) Resolution degradation by cross talk will be negligible because of its spectral sensitibity and stagger arrangement of sensor elements
(3) Atmosphere will have considerable effect on spatial resolution
(4) Spatial resolution in along track will be affected by the platform motion. It depends on the start point of the integration. Capability of integration time is usefull for improvement of the resolution in along track.

More detailed simulation using the real characteristics of AVNIR is now in the planning stage.

REFERENCES

H.Koshiishi,et al.: Simulator of pushbroom scan spectrometer, Proceedings of IGARS S'93, Aug.18-21, 1993, Tokyo, Japan, pp 882-884

H.Koshiishi : Why the pushbroom scan, and what the limitation ?, Proceedings of 18th ISTS, 1992, Kagoshima, Japan. pp 1925-1930

Prévision du rendement de l'AVNIR par simulateur

RÉSUMÉ

Traduction non disponible pour cause de livraison tardive du résumé définitif

VIRTUAL GIS, A NEW REALITY

By Nickolas L. Faust
Dharmajyoti Bhaumik
Ryan Woodard
Dung Vu

Georgia Tech Research Institute
Georgia Institute of Technology
Electromagnetics and Physical
Sciences Laboratory
Atlanta, Georgia 30332

KEY WORDS: GIS, Visualization, Remote, Sensing, Geographic, Virtual Reality Computing, Rendering

ABSTRACT

The merger of Geographic Information Systems (GIS) analysis with multidimensional visualization has allowed the development of a unique interactive environment for GIS systems. Instead of the traditional top down viewing of GIS layers and analyses using those layers, a three dimensional approach is described that combines GIS functionality with a fast 3 dimensional code developed at the Georgia Tech Research Institute (GTRI). While three dimensional perspective viewing has been common in flight simulation and has been used recently in display of remote sensing and GIS data sets, it has always been used primarily as a glitzy method to display basic data and pregenerated analyses. By combining fast rendering within a GIS structure, direct query of attributes can be added to visibility analysis allowing the practical use of perspective imaging as a direct interface into a GIS analysis system. This interface is a natural method for a user to interact with spatial data sets.

I. INTRODUCTION

Traditional methods of GIS display and analysis normally provide for the display of vector and raster GIS data in a top down, seen from above presentation. Overlays, proximity analyses, and other GIS analysis functions are always performed in a two dimensional view of the database. Many of these analyses include terrain elevation data or derived parameters such as slope and aspect. Functions such as terrain masking help the user understand the influence of the local terrain on visibility from a point within the database. A line of sight analysis will show a terrain profile between points of interest in the database. Visibility is becoming an increasingly important factor in GIS analysis, however, a two dimensional or one dimensional representation of the area that can be seen is not as effective as a full 3 dimensional view. A three dimensional perspective view shows not only the GIS layers of interest, but also what can be seen from a particular location within, above, or below the terrain.

A dynamic method for combining both the multi-layer weighting, proximity, and visibility operations as well as interactive query into a single, powerful, interface into GIS functionality is being developed at the Georgia Tech Research Institute (GTRI). This should not be viewed as the development of another new GIS system. Instead, it should be viewed as a synergistic combination of strengths of a rapidly expanding GIS field and the similarly expanding visualization environment. A GIS system has grown more and more sophisticated as new models, techniques, and applications are included. Combinations of raster and vector information is occurring in several commercial GIS systems bring further capabilities for display and analysis of spatial data.

During the same time, visualization techniques have come out of the closet of the computer graphics artists and have started to be used in practical GIS applications. Simulation techniques developed over a number of years by the flight simulation industry using specially designed hardware, are now becoming viable on general purpose hardware.

The synergistic combination of these two worlds into a truly interactive' virtual' GIS is now becoming a reality.

II. VISUALIZATION IN CURRENT GIS SYSTEMS

Current generation GIS systems generally view, analyze, and perform functions on spatial data (vector/raster) is a top down two dimensional manner. Complex analyses may be performed on the spatial data variables using weighting indexes, proximity functions, and a variety of multi-layer models. Each of the source layers of information as well as the analysis result is usually shown as seen from above. In many cases visibility of one part of the spatial data set from another position within the data set becomes as issue that must be considered. Example questions might involve whether a strip mine area can be seen from a major scenic highway, or whether a company's view from its building will be occluded by the proposed erection of another building between it and the ocean. Visibility is generally handled in one of two ways, terrain masking, and perspective viewing. For terrain masking, a number of line of sight calculations are made in all directions from the viewers position. A line of sight calculation normally involves the reading of the terrain and feature data along a particular direction from the viewer. If there are locations along this line of sight that are higher in elevation than other points along the line of sight but farther away, then some of the terrain will not be seen by the viewer. These seen and not seen areas in all directions create a terrain mask in which areas that can be seen from the viewer location are differentiated from areas that can not be seen. The areas that can not be seen are said to be masked by the terrain. This is a two dimensional representation of a visibility index.

Perspective viewing allows the calculation and display of exactly what will be seen from a particular viewpoint. The perspective presentation of spatial data mimics the way that the eye perceives what can be seen. A view can be calculated that presents a very natural and interpretable image to the GIS user. The image, depending on the quality of the input image and GIS data, may show exactly what the viewer will see from that location. GIS systems which include multispectral imagery along with topography and other GIS layers can be used to show for example, what can be seen from the top of Stone Mountain, near Atlanta if one looks to the north west. Images from satellite and aircraft can record information in the visible, (red, green, blue) parts of the electromagnetic spectrum. Thus, when these images are draped over the surrounding terrain, the view closely replicates the human's view from that location. Atmospheric haze can be added to the generated images, or one can look at the view as if no haze were present; for example, after a fall cold front passes through.

Perspective views have been generated for a number of years to show the view from a given geographic position. In general, however, perspective views are only used as a presentation method for final results and are not used in the analytical process itself. One of the principal reasons for this is the time it takes to generate a realistic perspective image. We will show in this paper that if this rendering can be done in near real time with modern technology, the perspective view will become the most natural interface for many GIS applications.

III. HIGH QUALITY/HIGH SPEED RENDERING

GTRI initially developed perspective scene rendering in the mid 1970's. Since that time development has occurred both in computing and graphics hardware , but also in the software implementation of perspective view algorithms. Whereas, most of the perspective generation techniques were implemented in somewhat system specific Fortran code in the 70's and early 80's, today's trend is toward C and C++ versions of rendering algorithms that can be directly implemented on almost any Unix workstation. There are many aspects of rendering that must be considered when designing or implementing a system. One must determine the necessary quality of a perspective image that is required by an application. Next, one must determine how fast that rendering process needs to happen to adequately support the application. Usually, the higher the quality of the rendered scene, the longer the time is for rendering. For a number of years, Georgia Tech has made the distinction that it would rather produce higher quality rendering than faster rendering of lower quality images. With today's technology, that distinction may not have to be made.

Flight simulation companies, until recently, have always relied on high speed specially designed hardware to achieve real time rates. When a scene was to be rendered, a multi-processing or parallel environment was used to render terrain and natural and manmade objects. Before the use of high speed random access memory (RAM) and video RAM, a display list of polygons and vectors was used to sort which things within the terrain and objects should be placed in front of others. Since this sorting and the display of the result had to be accomplished in 1/30th of a second, only limited resolution databases and object descriptions were allowed. In all except the very high end simulators (6 to 10 million dollars), the real time simulation looked cartoonish and was not totally effective in training. A person training on the system often would get distracted by graphical artifacts instead of concentrating on the content of the image presented. The larger simulation systems used a large number of specialized processors, and the computing engines, themselves, sometimes occupied whole rooms.

The advent of cheaper RAM and video RAM caused most manufacturers to switch to a more effective rendering technique using a depth buffer approach instead of sorting polygons and vectors by distance and then displaying the results. The depth buffer replaced the sorting function of the rendering. The depth buffer was memory associated with every image picture element (pixel) that held the distance between the viewer and that point on the terrain. Individual terrain polygons could be then rendered in a brute force manner using the depth buffer to automatically provide the sorting. As each polygon from the terrain or object was considered, the world coordinates of each vertex of the polygon was geometrically transformed into the output image space and an interpolation process was used to fill in the color values within the output image based on the distance in the depth buffer for that image pixel. By the late 1980's, the depth buffer approach was being used even in the most expensive simulation systems. To achieve the necessary update rates, parallel processing and multi-processing hardware were still used with the depth buffer common to all rendering processes.

Texture mapping was used in many of the simulation systems to provide artificial detail in areas of the terrain that were of little interest. Generic, mathematical textures were developed for tree canopy, grass, brick buildings, etc. where a regular function could add detail to a terrain polygon. Much expertise was required in a simulation design to determine which parts of a database needed high resolution, and which parts could get away with mathematical texture. Techniques such as Delauney triangulation were used to optimize the number of polygons necessary to adequately represent a given area's terrain. Areas that had a complex terrain required more and smaller polygons (generally triangles or rectangles) while areas that were flatter could be adequately represented by fewer, larger polygons. When using mathematical texture, this technique could result in a reasonable representation of the terrain, and more of the compute time could be spent on rendering high resolution versions of the natural and manmade objects.

Normally, a real time system can process a given number of polygons in 1/30th of a second. If more detail is given to the terrain, then less detail will be available to the object. The size of the individual polygon may not be a major time factor. This tradeoff between realism and speed has always existed in simulation.

In the late 1980's photo texture became the trend in simulation systems. Instead of generic or mathematical texture functions to determine the color of a specific polygon, an image of a real world object or terrain was used. When the terrain polygon was projected to the screen, an interpolation or mathematical function was not used to determine how to fill the projected polygon. Instead, the real world image was transformed through the same perspective transformation and pixel by pixel used to populate the polygon. Photo texture is especially effective when it is necessary to portray a very complex setting such as an urban landscape. Buildings within this urban environment might take very many polygons to describe in a way that allowed the viewer to see a realistic view. Instead, a simple rectangular polygon could be projected and a digitized photograph of a real building could be mapped onto that projected polygon. A simulation system that takes advantage of photo texture might spend less time on the polygon projection and concentrate on the relatively straightforward image processing functions to remap the image to the polygon.

The above techniques use a data base to screen type of rendering in which within a limited view area, all polygons are rendered using the depth buffer to mediate visibility. It can be seen that this technique is very time dependent on the size of the terrain and object database in terms of the number of polygons to be projected and filled. For databases using real imagery as the terrain texture information, it is not uncommon for over 100,000,000 polygons to be in a spatial database. For example, a merge of SPOT satellite information (60 km x 60 km) with a 10 meter pixel size with a portion of a Landsat Thematic Mapper scene (185 km x 185 km) with a 30 meter pixel size would easily give over 70,000,000 triangles to be rendered if the whole data base could be seen. Several other methods are often used in the generation of perspective images, ray tracing and inverse ray tracing.

Ray tracing is a very straightforward procedure which often takes significant computer resources, but that generally results in high quality rendered images. Ray tracing assumes that there is at least one light source radiating light rays onto the spatial database including terrain and objects. Parallel rays are cast from the source toward the terrain. As each ray intersects the terrain it is either absorbed, reflected, or transmitted through the terrain material. If it is reflected, it may intersect another part of the data base or it may be reflected away from the data base. If it intersects another part of the data base, another calculation of absorption, reflection, or transmission must be performed. Ray tracing usually limits the number of multiple bounces that are performed.

Those rays which are reflected toward the viewer's eye represent the image that is generated for a specific user view pyramid. The ray tracing technique must be able to decide which terrain or object polygons are intersected by initial rays or reflected rays. The search through all possible polygons for each ray is very time consuming. Many optimizations are used to reduce the number of intersections that must be calculated.

A ray traced image may show very realistic effects such as reflections off of shiny surfaces, and generally is the type of image that is used in movie sequences.

Inverse ray tracing is also popular in rendering. In this case instead of tracing a ray from the source until it intersects the terrain and finds its way into the view pyramid, the ray is cast from the viewer's eye, through image pixels in the imaging plane, and onto the terrain. This technique must also look for the intersection of the ray with the terrain polygons, but much fewer rays have to be cast. The terrain reflectance properties are assumed to be lambertian, in which for a given source ray, reflection occurs in all directions. Multi-bounce scattering is not easily accommodated within inverse ray tracing.

IV. GTRI RENDERING APPROACH

GTRI has developed a generic C rendering approach using the methodology of data base to screen rendering as described above. This technique implements photo texture mapping for both terrain and objects, can handle arbitrarily large spatial databases, incorporates antialiasing techniques for motion rendering (simulation), and runs in near real time (approximately 10 seconds) on a Sun SPARC 1 class workstation. Dynamic indexing into multiple levels of database resolution creates appropriate background fuzziness and antialiasing while reducing the total number of polygons to process. Image data or GIS feature layers may be draped over the terrain, giving a user the capability to dynamically see raw or analyzed information in a natural perspective. Using a graphical user interface (GUI) a user may navigate throughout the database interactively, evaluating dynamically the visibility of certain features from his autonomous path. The perspective images of the spatial data set provide the ability for the user of the system to totally immerse himself within the spatial data base with ultimate free movement. The scenes that are created as the user moves are high quality, detailed, views of the spatial data set. The user is truly immersed into a virtual spatial data set using a Virtual GIS (VGIS).

As the speed of rendering and display increases with the Joy's Law increase in computing power , and as more of the process is implemented in parallel, full real time immersion is possible. While virtual immersion within the spatial data set is practical in the near term (1 to 2 years), the speeds for typical GIS spatial analysis functions should also be increasing. Thus, a user might place himself on a hillside and watch a rainstorm drop water into a watershed (a graphical illustration of a meteorology model), watch as the water finds its way into streams (as modeled by watershed runoff models), and see the river rise and create flood hazards. This is the very beginning for the merger of modeling, GIS, and visualization.

V. REALTIME GIS INTERFACE

In addition to the ability to immerse oneself within spatial data, a user must also be able to perform the natural GIS query functions and dynamically see the results of GIS analysis. The query function is implemented with the dynamic movement control structure of the view model. All interaction occurs with the rendered image. The user can place himself high above the database and essentially see a two dimensional view of the spatial data, or he can zoom into the data to view it from a specified view point. At any time, the user may query any point within the perspective image and retrieve geographic coordinates, database coordinates, the values of the color in rgb, and the values or attributes of associated GIS layers at that exact same location. When multiple GIS raster layers are available, the user can select either the original true color image, the original GIS layers (such as landcover), or analysis layers such as wood duck habitat to drape over the terrain. The user can then check the validity of the layers by the way they correspond to the landscape form. For example, Landsat image data in true color may be overlaid with Corps of Engineers 100 year flood plan information with the result being shown in perspective. If the flood plain does not correspond to topographic lows near the river, then the observer will immediately pick out the conflict.

VI. THE OLYMPIC VILLAGE TEST CASE

Georgia Tech assisted the Atlanta Organizing Committee in winning the 1996 Olympic games for the city of Atlanta. The campus of Georgia Tech will become the Olympic Village, and many activities are now being planned for the campus. GTRI has constructed a high resolution geographic data set based on true color aerial photography, 2 foot contour elevation data from planning maps, and spatial data sets for building, parking lot, road, and utilities locations. These data have been supplemented by attribute information for all major buildings on the campus. For each building, attributes were stored showing the name of the building, the use of the building, the height of the building, and the number of floors in the building. Roads were associated with road names. In addition to the GIS attribute information, the footprint of each building was located as well as the location and size of every tree on the campus. For major buildings on the Tech campus, photographs were take of as many sides of the building as could be seen without major vegetation obscuration. Buildings other than the major buildings were given a generic brick texture which is characteristic of many buildings on the campus.

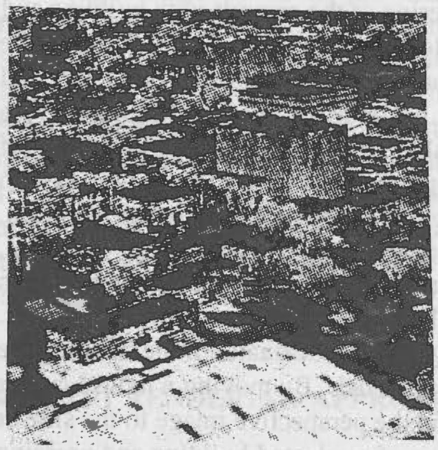
The original data base of the Georgia Tech campus is being extended towards Peachtree Street on the east and downtown on the south. Similar photographs and information will be gathered on most of the buildings by graduate city planning students. After the individual building photographs have been taken, and correlated with the GIS attributes given above, the images are scanned into digital form and processed to remove optical and perspective distortion.

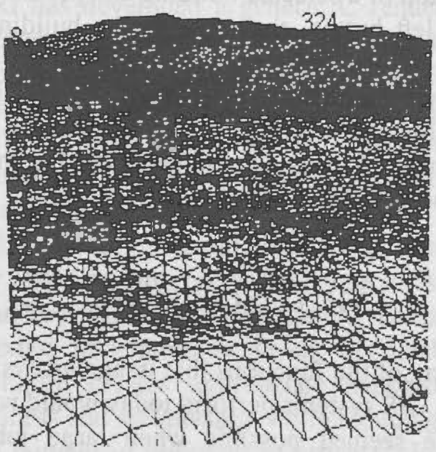
Figure 1 shows the Georgia Tech campus in a perspective view with no objects displayed. Figure 2 shows a wireframe rendering of the campus with building and tree locations shown. Figure 3 shows the campus with phototexture buildings and trees. The black lines on figure 3 are a crosshair cursor. The College of Computing was identified by the system as being located at the crosshair position. Figure 4 gives the name of the building, its primary use, the number of floors in the building, and the building height. Geographic coordinates of the building are also reported. By moving a mouse cursor within the perspective image to the side or the top of a selected building and pressing the left button, the name of the building and all its attributes are displayed on a popup information screen. This interaction is dynamic so attributes of a number of buildings can be queried at the same session.

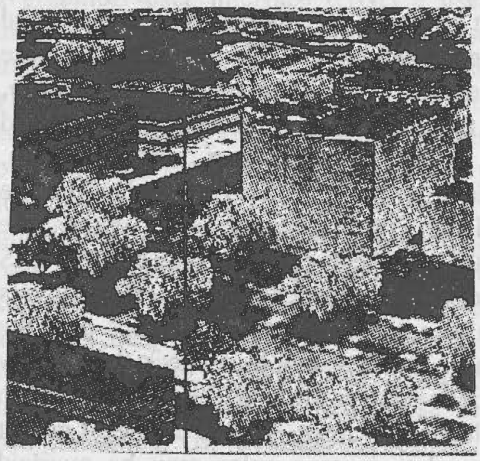
VII. CONCLUSIONS

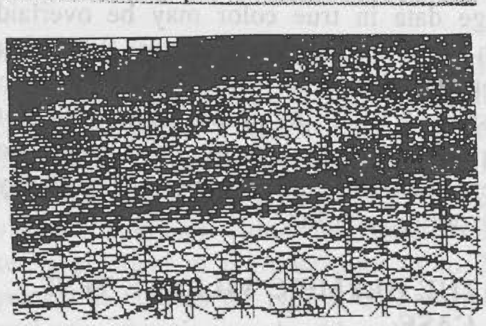
Many of the functions that image analysis and GIS systems perform today would not have seemed possible when large jobs were submitted to mainframes for overnight computer runs. We used to justify the overnight turnaround by saying that it gave one time to think about the changes that he might want to make in the process. Now, we interactively work with workstations and dynamically change programs for immediate execution. The sophisticated and unsophisticated analyses being performed in short timeframes would have taken years to run just 15 years ago.

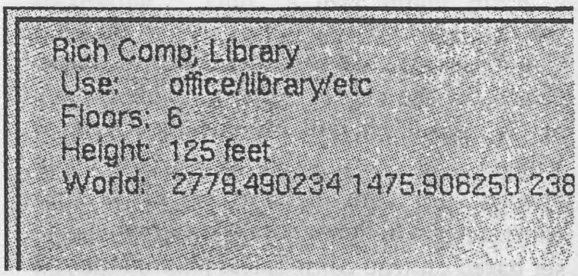
The advances in the computing industry can be combined with human interaction methodologies of visualization and the spatial data management capabilities of a GIS to give a dynamic and virtual 3 D spatial data environment. As this environment becomes more and more a real time presentation, display, and interaction capability, our whole approach to spatial analysis may change. The potential for the integration of GIS with a virtual reality interface is almost here.











SIG virtuel à trois dimensions

RÉSUMÉ

La technologie des systèmes d'information géographique (SIG) se répand maintenant dans le monde entier comme une traînée de poudre. Étant donné que la gestion des données spatiales devient par ailleurs un outil essentiel dans un nombre sans cesse croissant de disciplines, la masse même de ces données, en croissance exponentielle, finira par constituer un problème majeur, alors que les cartes numériques tendent à remplacer les cartes analogiques traditionnelles sur papier et que les données d'images sont couramment utilisées dans diverses applications comme par exemple l'éditique.

Le présent document porte sur les diverses fonctions et les complexités des SIG et sur la manière dont les technologies naissantes de visualisation et de réalité virtuelle peuvent redéfinir un nouveau mode d'interaction homme-machine en ce qui a trait aux SIG.

Avec le perfectionnement des nouvelles techniques de rendu et les vitesses croissantes de traitement qu'autorisent les circuits intégrés des ordinateurs à usage général, la mise au point d'un éventuel SIG virtuel et vraiment interactif est dorénavant envisageable. Un tel SIG pourra intégrer toutes les fonctions traditionnelles et se prêterait tout naturellement à l'analyse par couches superposables et à la visualisation en perspective. L'utilisateur pourra ainsi interroger le système de façon aussi spontanée qu'il peut s'interroger lui-même devant certains éléments du décor naturel, en demandant par exemple à qui appartient ce terrain là-bas au sommet de la colline. La capacité du système à répondre à de telles interrogations permettra une utilisation beaucoup plus intuitive et aisée des énormes possibilités des SIG évolués, par rapport à celles de leurs prédécesseurs. Rappelons que les données d'image constituent une caractéristique importante des SIG virtuels, la perception de l'information étant plus aisée lorsqu'elle se fait visuellement, à partir d'un point de vue particulier; l'observateur de l'image virtuelle aura également plus de facilité à faire le rapprochement entre cette image et l'analyse de l'information SIG dont il a besoin.

UN LANGAGE DE REQUETES A OBJET POUR LES IMAGES

Mohamed EL ANSARI et Liming CHEN

Laboratoire HEUDIASYC ura cnrs 817, Centre de recherche de Royalieu, Université de technologie de Compiègne, B.P. 649 - 60200 Compiègne Cedex, France.

Les applications des bases de données deviennent de plus en plus diversifiées. Les systèmes de bases de données actuellement commercialisés et développés sur la base de la gestion des données formatés sont de moins en moins efficace pour ces applications. Nous nous sommes intéressé aux bases de données image. Les données (images digitales ou graphiques) sont organisés sous forme d'une hiérarchie d'héritage. Les opérateurs sur les images ont été identifié et associé aux différentes classes d'image. Nous proposons un langage de requête de type SQL orienté objet pour la gestion de ces bases de données. Il utilise les classes et les méthodes définies dans la hiérarchie d'héritage. Le langage combiné avec la hiérarchie d'héritage offre un environnement puissant pour supporter n'importe quel type d'application image.

1. INTRODUCTION

Les bases de données orientées objet ont pour ambition de satisfaire des applications non traditionnelles qui sont les applications de CAO, de multimédia, de cartographie, de bureautiques, etc. Dans ces nouvelles applications, les données sont non seulement volumineuses mais également complexes; elles peuvent être des textes, des graphiques, des images ou du son.
Nous nous intéressons dans ce papier aux bases de données d'images. Les informations à gérer dans ce type de base peuvent être classées en cinq catégories (Grosky, 1989):

- Données iconiques :
  Les images elles-mêmes sont stockées sous forme analogique ou digitale.
- Données sur les images :
  Ces données peuvent être la résolution d'une image ou divers formats de descriptions.
- Informations extraites des images :
  Ces informations proviennent des traitements d'images. Elles contiennent les structures topologiques/numériques, les composantes structurelles d'une image et les correspondances entre ces composantes.
- Relation monde-image :
  Ce sont les relations entre les composantes d'une image et les entités du monde réel.
- Données sur le monde réel :
  Ce sont les données textuelles classiques décrivant le monde réel.

Un système de base de données d'images doit faciliter le stockage et la gestion de ces cinq types d'informations. Ces systèmes doivent être capables de traiter des requêtes sur des données iconiques basé sur leur contenue et de rechercher des informations sur des données textuelles et iconiques à partir d'autres données iconiques ou textuelles.
Nous proposons dans ce papier un langage de requête pour la manipulation des bases de données d'images. L'originalité de notre approche est la structuration des données d'images selon le principe orientée objet. En effet, les données d'images sont classées et organisées dans une hiérarchie d'héritage. Les opérateurs les plus utilisés ont été identifiés et associés à différentes classes. Le langage est du style SQL et fait appel aux classes et méthodes définies dans la hiérarchie. Ainsi, la hiérarchie d'héritage de données d'images combinée avec le langage fournit un environnement puissant pour tout type d'applications d'image.
La suite du papier est organisé comme suit. Dans le paragraphe 2, on donne une brève présentation sur le gérant d'objets VROOM sur lequel le langage sera implanté. Le paragraphe 3 sera consacré à la description de la hiérarchie d'héritage d'images et des différentes méthodes identifiées. Nous présenterons ensuite au paragraphe 4 le langage permettant la manipulation des données d'images. Le paragraphe 5 contient une comparaison de notre approche avec d'autres propositions et la conclusion.

2. LE GERANT D'OBJET VROOM

VROOM est un gérant d'objets développé par l'équipe Rapid du laboratoire MASI à l'université Paris VI. VROOM est un système composé de plusieurs couches. Le noyau du système contient les mécanismes de base permettant la création, la manipulation et la destruction d'objets non structurés. Le noyau gère les accès concurrents, la persistance et les transactions atomiques sur ces objets. Au dessus de ce noyau, une couche est construite pour fournir des fonctionnalités riches permettant par exemple la structuration des objets, le groupement des objets (clustering) ou la création des indexes. L'interface utilisateur du système permet la création et la manipulation d'objets similaires aux objets C++ avec quelques contraintes.

2.1 L'architecture du système

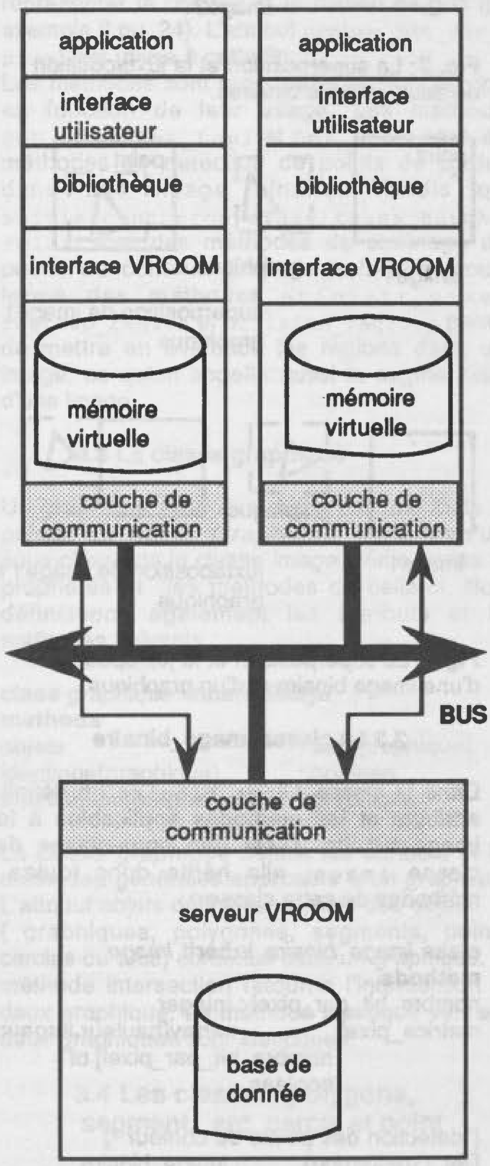


Fig. 1 : L'architecture de VROOM

VROOM a une architecture clients/serveur (fig.1). Les applications tournent sur les stations de travail et utilisent une bibliothèque qui gère notamment une mémoire virtuelle dans la station. L'interface VROOM permet la manipulation des objets par les applications. Le serveur pour sa part gère le stockage des objets, la concurrence et la consistance entre les différentes applications accédant au serveur.

2.2 L'interface utilisateur du système (VROOM/C++)

L'interface utilisateur de VROOM permet aux programmeurs de développer des applications C++. La persistance des objets est assuré par le système et des extensions ont été ajoutées au langage C++ pour qu'il soit adapté aux applications bases de données. Les extensions les plus importantes sont : la gestion des transactions, les collections et le groupement des objets (clustering). Un objet est définie de la même façon que dans C++ mais son identité n'est pas définie par son adresse physique mais une adresse logique. VROOM fournit quatre types de collection : les séquences, les listes, les ensembles et les bags. L'héritage est défini de la même façon que dans C++ .

3. HIERARCHIE DES DONNEES IMAGE

Les données images sont structurées en classes. Les classes sont organisées en une hiérarchie d'héritage (fig. 2). La classe à la racine de cette hiérarchie est appelée *image*. Elle se spécialise en deux sous-classes : *image-binaire* et *graphique*. La sous-classe *graphique* se spécialise aussi en des sous-classes : *polygone, cercle, segment, arc* et *point.*

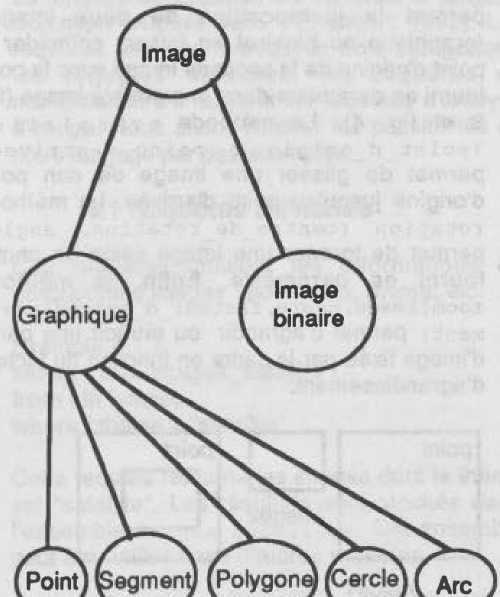


Fig. 2: La hiérarchie d'heritage des données d'image

3.1 La classe image

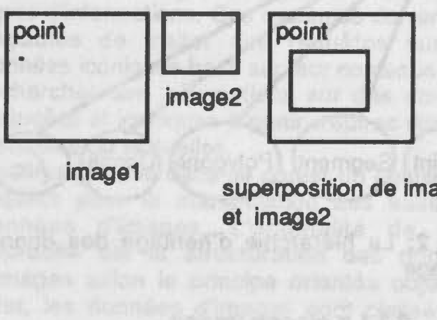
Les propriétés et les méthodes qu'on associe pour la classe image doivent pouvoir s'appliquer à tout objet image, qu'il soit binaire ou graphique. On a ainsi identifié les propriétés ou méthodes suivantes:

class image
methods

titre	: string
date	: Date
auteur	: Personne
theme	: string
origine	: Point
hauteur	: int
largeur	: int
superposition(image,point)	: image
juxtaposition (image, point)	: image
translation(point_d_origine, point_d_arrivee)	: image
rotation(centre_de_rotation, angle)	: image
zoom(image,cadre, facteur_d_agrandissement)	: image

Une image est donc définie par son titre, la date de la création, son auteur, son thème, son point d'origine qui est le coin supérieur gauche, sa hauteur et de sa largeur.

La méthode `superposition(image, point)` permet la superposition de deux images (graphique ou binaire) en faisant concider le point d'origine de la seconde image avec le point fourni en paramètre dans la première image. La méthode `juxtaposition(image, point)` permet la juxtaposition de deux images (graphique ou binaire) en faisant coïncider le point d'origine de la seconde image avec le point fourni en paramètre dans la première image (fig. 3 et fig. 4). La méthode `translation (point_d_origine, point_d_arrivee)` permet de glisser une image de son point d'origine jusqu'au point d'arrivée. La méthode `rotation (centre_de_rotation, angle)` permet de tourner une image selon un centre fourni en paramètre. Enfin, la méthode `zoom(image,cadre,facteur_d agrandissement)` permet d'agrandir ou rétrécir une partie d'image fixée par le cadre en fonction du facteur d'agrandissement.



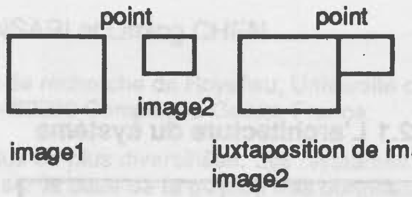


Fig. 3: La superposition et la juxtaposition de deux images binaires.

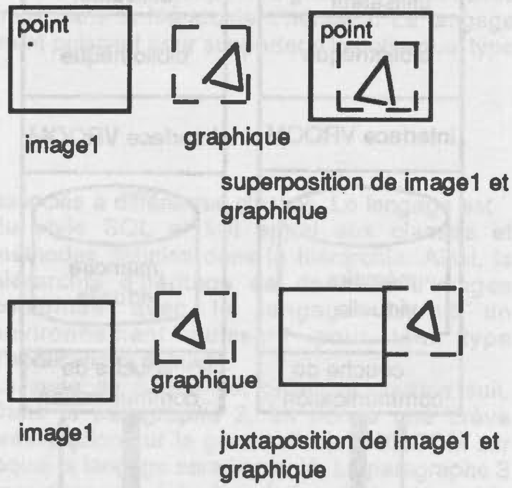


Fig 4: La superposition et la juxtaposition d'une image binaire et d'un graphique

3.2 La classe image_binaire

Dans la classe `image_binaire` on définit les attributs et les méthodes applicables à toute image binaire. C'est une sous-classe de la classe `image`, elle hérite donc toutes les méthodes de cette classe :

class *image_binaire* **inherit** *image*
methods

nombre_bit_par_pixel	: integer
matrice_pixel	: array[hauteur,largeur, nombre_bit_par_pixel] of boolean
{*detection des points de contour *}	
Det_Grad(string)	: image_binaire
Det_Lapl(string)	: image_binaire
Det_Comp(string)	: image_binaire
{* chaînage des points de contour*}	
Suivie_Cont()	: graphique
Prog_Dynam()	: graphique
Trans_Hough()	: graphique
Relax()	: graphique
{*segmentation de l'image *}	
Etiquettage_pixel()	: graphique
Regroup_region()	: graphique
Division_regroup()	: graphique

Une image binaire consiste en une matrice de pixels. Chaque pixel indique la couleur ou le niveau de gris d'une portion de l'image. Un pixel peut être codé par un simple bit pour représenter le noir et le blanc ou par plusieurs bits pour représenter la couleur et le niveau de gris (par exemple 8 ou 24). L'attribut `nombre_bit_par_pixel` est utilisé à cette fin.

Les méthodes sont organisées en trois groupes en fonction de leur usage. Les méthodes `Det_Grad`, `Det_Lapl` et `Det_Comp` sont des méthodes de détection de points de contour dans une image binaire. Tandis que `suivie_cont`, `prog_dynam`, `trans_hough` et `relax` sont des méthodes de chaînage des points de contour. Enfin le troisième groupe formé des méthodes `etiquett_pixel`, `regroup_region` et `division_regroup` permet de mettre en évidence les régions dans une image, ce qu'on appelle aussi la segmentation d'une image.

3.3 La classe graphique

Un graphique peut être défini par une suite de points. La classe *Graphique*, en tant qu'une sous-classe de la classe *image*, hérite toutes les propriétés et les méthodes de celle-ci. Nous définissons également les attributs et les méthodes suivants :

class *graphique* **inherit** *image*
methods
objets : set (graphique)
identique(graphique) : boolean
intersection(graphique) : graphique

La classe *graphique* définie les attributs et les méthodes générales applicable à un graphique. L'attribut objets définit l'ensemble des objets (graphiques, polygones, segments, points, cercles ou arcs) contenus dans un graphique. La méthode intersection retourne l'intersection de deux graphique. La méthode identique vérifie si deux graphiques sont identiques.

3.4 Les classes: polygone, segment, arc, cercle et point

class *polygone* **inherit** *graphique*
methods
contour : set (segment)
P_perimetre() : int
P_surface() : int

L'attribut `contour` contient les segments définissant le polygone. Les méthodes `P_perimetre` et `P_surface` définissent le périmètre et la surface d'un polygone .

class *segment* **inherit** *graphique*
methods
point_debut : point
point_fin : point
centre() : point
longueur() : int
pente() : string

`point_deb` et `point_fin` sont les points extrêmes du segment. Les méthodes `Centre`, `longueur` et `pente` définissent respectivement le centre, la longueur et la pente du segment . Nous définissons ci-dessous les classes `cercle`, `arc` et `point`. Nous ne donnons pas plus d'explication car leur signification semble évidente.

class *cercle* **inherit** *graphique*
methods
centre : point
rayon : int
C_surface() : int
C_perimetre(): int

class *arc* **inherit** *graphique*
methods
point_deb : point
point_fin : point
point_arc : set (point)

class *point* **inherit** *graphique*
methods
abscisse : int
ordonnée : int

4. LE LANGAGE DE REQUETES

Le langage de requêtes est basé sur le langage SQLobjet proposé dans (El ansari, 1992). Les requêtes sont regroupées en trois catégories : les requêtes utilitaires, les requêtes de manipulation d'images et les requêtes d'analyse d'image. Nous allons illustrer les possibilités de notre langage par des exemples.

4.1 Requêtes utilitaires

Ces requêtes utilisent les informations de description d'images : nom, date, thème, etc.

Exemple 4.1.1.
select I **into** images_satellite
from I **in** images
where I.thème = "satellite"

Cette requête retourne les images dont le thème est "satellite". Les résultats sont stockés dans l'ensemble `images_satellite`. Cet ensemble peut être utilisé dans d'autres requêtes.

Exemple 4.1.2.
select I **into** images_90_s
from I **in** images
where I.date.superieur(Date(1,1, 1990))

La requête retourne les images dont la date est supérieur au 1er janvier 1990. Le résultat est stocké dans l'ensemble images_90_s.

Exemple 4.1.3.
select I.hauteur, I.largeur, I.bit_par_pixel
from I **in** images_binaires
where I.nom = "toto"

Cette requête retourne la hauteur, la largeur et le nombre de bits par pixel de l'image "toto".

Exemple 4.1.4.
update I (nom = "machin", date = Date(10,11,1991))
from I **in** images
where I.nom = "toto"

Cette requête modifie le nom et la date de l'image "toto".

4.2 Opérations de manipulation d'image

Les opérations de manipulation d'images accomplissent certaines transformations sur les images pour en fournir différentes perspectives.

4.2.1. Opérations de zoom : Ces opérations permettent d'agrandir une image ou un objet à l'intérieur d'une image.

Exemple 4.2.1.
select I.zoom(2)
from I **in** images
where I.nom = "toto"

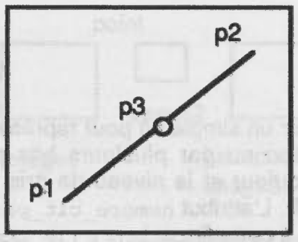
Cette requête retourne l'image "toto" agrandi deux fois.

Exemple 4.2.2.
select I **with** J.zoom(3)
from I **in** graphique, J **in** I.objets
where I.nom = "toto" **and**
J.cente.egal(point(30,30))

La requête retourne le graphique "toto" avec le cercle, dont le centre est le point (30,30), agrandi deux fois.

4.2.2. Opérations de rotation : Ces opérations permettent la rotation d'une image ou d'un objet à l'intérieur d'une image d'un certain angle autour d'un axe.

Exemple 4.2.3.
select I **with** J.rotation(10, point(8,8))
from I **in** graphique, J **in** I.objets
where I.nom = "toto" **and**
J.point_deb.identique(point(4,20))
and J.point_fin.identique(point(20,4))



toto

La requête retourne l'image toto avec la droite définie par les deux points `p1(4,20)` et `p2(20,4)` retournée autour du p3(8,8) d'un angle de 10 degrés.

4.2.3. Opérations de superposition et de juxtaposition : Ces opérations permettent la superposition ou la juxtaposition de deux images.

Exemple 4.2.4.
select I.superposition(J,I.origine())
from I **in** images_binaire, J **in** images_binaire
where I.nom = "toto" **and** J.nom = "machin"

Cette requête retourne l'image J superposée sur l'image I en faisant coincider le coin supérieur gauche des deux images.

Exemple 4.2.5.
select I.juxtaposition(J,P)
from I **in** graphique, J **in** image_binaire, P **in** point
where (I.nom = "graph1") **and** (J.nom = "bin1")
and (P.abscisse = I.origine.abscisse + I.largeur)
and (P.ordonnée = I.origine.ordonnée)

Cette requête retourne le graphique `graph1` juxtaposé avec l'image binaire `bin1`, en faisant coincider le coin supérieur droit du graphique `graph1` avec le coin supérieur gauche de l'image binaire `bin1`

4.3. Opérations d'analyse d'images binaires : Ces opérations traitent des images binaires pour en extraire des primitives.

4.3.1. Détection des points de contour: Ces opérateurs sont appliquées à une image binaire pour détecter les points de contour des objets contenue dans cet image. Le principe repose sur l'utilisation des masques (Dana, 1982).

Exemple 4.3.1.
select I.Det_Lapl(4-laplacien) **into** images_bin_lapl
from I **in** image_binaire
where I.nom = "toto"

Cette requête traite l'image binaire `toto` par la méthode de détection des point de contour du laplacien (le masque utilisé est le 4-laplacien). Le résultat est une image binaire. Elle est stockées dans l'ensemble `images_bin_lapl` qui contient les images binaires traitées par la méthode du

laplacien .
Les détecteurs et les masques qui peuvent être utilisés sont :
- `Det_Grad(masque)` : détecteur du gradient. Masque : Prewitt, Roberts, Sobel ou Isotropique.
- `Det_Lapl(masque)` : détecteur du Laplacien. Masque : 4-Laplacien ou 8-Laplacien.
- `Det_Compass(masque)` : détecteur de Compass. Masque: Prewitt, Kirsh, 3-niveau ou 5-niveau.

4.3.1. Chaînage des points de contour: Les contours s'obtiennent en reliant les points de contour extraits par l'une des méthodes précédentes.

Exemple 4.3.1.
Select I.suivi_contour() **into** images_suiv_cont
from I **in** images_binaire
where I.nom = "toto"

Cette requête obtient les points de contour de l'image `toto` par la méthode du chaînage du suivi de contour. Le résultat est un graphique stocké dans l'ensemble `images_suiv_cont` contenant les graphiques des images binaires traitées par cette méthode . On peut aussi faire appel aux autres méthodes :
- `Prog-Dynam()` : méthode de la programmation dynamique
- `Trans-Hough()` : méthode de la transformée de Hough
- `Relax()` : méthode de relaxation

4.3.2. Segmentation d'image : Ces opérations ont pour but de mettre en évidence des régions homogènes dans une image .

Exemple 4.3.2.
select I.Etiquett_pixel() **into** images_etiq
from I **in** images_binaire
where I.nom = "toto"

Cette requête traite l'image binaire `toto` par la méthode d'étiquettage de pixel. Le résultat est stocké dans l'ensemble `images_etiq` contenant les images binaires traitées par cette méthode.

5. COMPARAISON ET CONCLUSION

L'intégration des images dans les systèmes de bases de données a commencé avec les systèmes qui stockaient simplement les images dans des fichiers (EIDES[4] et IMDB[5]). Ces systèmes n'offraient pas de modèle de données, mais seulement un ensemble d'opérations (routines) pour accéder et manipuler les fichiers image. D'autres systèmes stockaient les informations sur les images dans une base de données relationnelle et les liaient aux images brutes stockées dans des fichiers (REDI/IMAID[4,6] et GRAIN[7,8]). Ils utilisaient des relations spéciales où chaque tuple représentait une image. Les opérations d'affichage et d'édition peuvent être appliqué aux tuples de ces relations. Tang[9] a proposé plus tard de représenter conceptuellement les images bruts dans le modèle de données sous forme de valeurs d'attribut. Les images restent stockées dans des fichiers, mais elle sont accessibles à travers le langage de requêtes. Une approche plus récente, PSQL[10], utilise le concept de type de données abstrait [11] pour représenter les images. PSQL supporte des domaines de description de données géométriques (points, segments et régions) définies sous forme de type données abstrait. Pour chaque type de données abstrait est définie un ensemble d'opérateurs. Cependant dans tous ces systèmes les descriptions de données et les opérateurs sont définies pour gérer un type d'image particulier (images satellite, Images géographiques, etc).
Nous avons proposé dans ce papier une alternative pour résoudre ce problème en utilisant un modèle orienté objet et en développant des fonctionnalités de base qui peuvent être utilisé dans n'importe quel type d'application image.

REFERENCES

Assman, K., R. Venema and K.H. Hohne, 1986.
The ISQL Language: A software Tool for the developement of pictorial information system in medicine. S-K. Chang, T. Ichikawa and P. Ligomenides (Eds.), visual languages, Plenum Press. p. 261-284.

Bordogna, G., I. Gagliardi, D. Merelli, P. Mussion, M. Padula,M. Protti, 1989.
Iconic Queries on Pictorial data. In Proc. IEEE whorkshop on visual langage, Rome. p. 38-42.

Brolio,J., B.A Draper, J.R Beveridge and A.R. Hanson, 1989.
ISR: A database for symbolic processing in computer vision. Computer,vol. 22, no 12. p. 22-32.

Chang, S.K., J. Reuss and B.H MacCornick, 1977.
An Integrated Relational Database System for Pictures. In Proc. IEEE Workshop on Picture Data Description and Management Chicago. IEEE Computer Society, catalog no. 77 CH 1187-4C. p 49-60.

Chang, N.S, and K.S. Fu, 1980.
Query-by-Pictorial-Example.
IEEE transaction on software enginneering, vol. SE 6, no 6.

Chang, N.S and K.S. Fu,1981
Picture Query Langages for Pictorial Information Systems. IEEE Computer, Vol. 14, no 11. p 23-33.

Chang, S.K, E. Jungert and Y. Li , 1989.
The Design of Pictorial databases based upon the theory of symbolic Projections. Proc. conférence on very large spatial databases, Springer-vergla.

Chock, M., A.F. Cardenas and A. Klinger, 1984.
Database structure and manipulation capabilities of a picture database management system (PICDMS). IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-6. p 484-492.

Dana, H.B, C.M. Brown, 1982.
Computer Vision. Prentice-Hall, INC. Englewood Cliffs, New Jersey.

EL Ansari, M. 1992.
Spécification d'un langage de requêtes pour une base de données orientée objet. Rapport de DEA, Université de technologie de compiègne, France.

Goodman, A.M., R.M. Haralick and L.G Shapiro, 1989.
Knowledge-Based computer Vision - Integrated Programming Langage and Data Management System Design. Compter, vol. 22, no 123. p 43-58.

Grosky, W.I and R. Mehortra, 1989.
Image Database Management. Special Issue of IEEE Computer, vol. 22.

Jagadish , H.V. and L.O Gorman, 1989.
An object model for image Recognition. Computer, vol. 22, no 12.

Joseph, T and A.F. Cardenas, 1988.
PICQUERY : A high Level query langage for pictorial database management. IEEE transaction on software enginneering, vol.14, no 5.

Kasturi, R. and J. Alemany, 1988.
Information Extraction from image Paper Based. IEEE transaction on software enginneering, vol.14, no 5.

Lee, Y.C and K.S. Fu, 1985.
A CAD/CAM Databases Management system and its Query langages. In langages for automation, S.K. Chang (Eds.), Plenum Pub. Co.

Lin, B.S. and S.K. Chang, S-K, 1980.
GRAIN-A Pictorial Database Interface. In Proc. IEEE Workshop on Picture Data Description and Management, Asilomar. IEEE Computer Society, catalog no. 80 CH 1530-5. p 83-88.

Lorie R.A, and A. Meier, 1984.
Using a Relational DBMS for Geographical Databases. Geo-processing, vol.2. p 243-257.

Meyer-Wergener, K, V.Y. LUM, and C.T WU, 1989.
Image Management in multimedia database system. In visual database systems, T.L Kunii (Eds). Elsevier Science Publishers B.V, North Holland.

Roussopoulos, N., C. Faloutsos and T. Sellis, 1988.
An efficient pictorial database system for PSQL. IEEE Transaction on software enginnering, vol. 14, no 5.

Samet, H., A. Rosenfeld, C.A. Shaffer, and R.E. Webber, 1984.
A Geographic Information System using Quadtrees. Pattern Recognition, vol. 17. p 647-656.

Tang, G.Y, 1980.
A Logical Data Organization for the Integrated Database of Pictures and alphanumerical Data. In Proc. IEEE Workshop on Picture Data Description and Management, Asilomar. IEEE Computer Society, catalog no. 80 CH 1530-5. p 158-166.

Woelk, D. and W. Kim, 1987.
Multimedia Information Management in an object oriented database system. Proc. 13th Very Large Database Conference. pp. 319-329 .

AN OBJECT ORIENTED QUERY-LANGAGE FOR IMAGES

Mohamed EL ANSARI and Liming CHEN
Laboratoire HEUDIASYC ara cnrs 817
Centre de recherche de Royalieu
Université de technologie de Compiègne
P.B. 649 - 60200 Compiègne Cedex
tél : (33) 44 23 44 85 - fax : (33) 44 20 48 13
e-mail : lchen@hds.univ-compiegne.fr

Abstract :

As databases applications become more and more diversified, the capabilities of the current commercial database management systems developed on the basis of handling formatted data become less and less satisfactory. We are interested in image databases. We propose a query langage for image databases. The image data (bitmap or graphics) are organized in hierarchie of inheritance. Operators on image are identified and associated with differents classes of image. The query langage is an object oriented SQL. It uses classes and methods defined in the hierarchie of inheritance. The langage combined with the hierarchie of inheritance provide a powerful environement for any kind of image application.

EVALUATION OF SOFTCOPY PHOTOGRAMMETRIC SYSTEMS: CONCEPTS, TESTING STRATEGY, AND PRELIMINARY RESULTS

Raad A. Saleh and Frank L. Scarpace
University of Wisconsin-Madison
ERSC, 1225 W. Dayton St., Madison, WI 53706, USA
Phone: (608)262-1585 Fax: (608)262-5964 Internet: raad@cae.wisc.edu, scarpace@macc.wisc.edu

Intercommission WG II/III

KEY WORDS: Digital, Softcopy, Photogrammetric Systems, Production, Evaluation, Testing

ABSTRACT

The field of photogrammetry is undergoing tremendous changes attributed primarily to the way data are handled, analyzed, and presented. The development of many softcopy photogrammetric systems is a characteristic of the direction of these changes. A significant segment of the user community is still cautious about these new systems coming to the marketplace with very impressive claims. Furthermore, it may be safe to assume that established photogrammetric mapping production facilities have already invested a considerable resources in plotting machines, photographic equipment, and operator training. Therefore, it would be it difficult to shift into a new direction, which not only will it render the existing equipment outdated, but it will also require new major acquisition. It is evident that no standardized evaluation yet exists that compares some softcopy systems to a conventional photogrammetric workstation. Such unbiased evaluation is needed to show users the strengths, and of course the weaknesses, of softcopy photogrammetric systems. The Laboratory for Softcopy Photogrammetry at the University of Wisconsin-Madison is undertaking an evaluation project of a number of commercially available softcopy photogrammetric systems. The selection of these systems is controlled by specific factors, which are budgetary constraints, limited resources of qualified testing staff, and willingness of manufacturer's participation. In this paper, evaluation objectives, strategy, and preliminary results are presented. In view of certain applications, softcopy systems are found to be competitive, if not superior, to conventional photogrammetric operations. A number of specific factors, however, must be carefully considered, such as data volume, operator skills, and automation.

1. INTRODUCTION

Softcopy photogrammetric systems are increasingly available in the commercial market, ranging in capabilities, performance, and cost. There is no doubt that these systems have the potential to provide low cost sound photogrammetric production. However, this evolving technology is by no means fully matured, at least for the present time.

The Wisconsin Department of Transportation (WisDOT) is currently considering options regarding the introduction of softcopy photogrammetry technology into its mapping production operations. While it is a given that conversion is eventual, the issue at hand is whether current softcopy technology conform to WisDOT standards of performance pertaining to mapping operations. There are several issues need to be addressed in answering this question. For this purpose, two high-end systems have been examined. These are the Intergraph ISPN and the

Helava SOCET SET. Other systems are being considered for testing as well.

This paper summarizes preliminary results of investigating the utility of using softcopy photogrammetric systems in place of selected WisDOT mapping operations. The evaluation concept, strategy, and results can be generalized for other agencies of similar production requirements.

2. CONCEPT AND STRATEGY

The evaluation project is designed to address:

a) Functions and Operational Characteristics: These include software, such as input/output, orientation procedures, and extraction; hardware, such as mono and stereo viewing, floating mark operations and coordinate measurements, memory, and storage capacity; and other system related operational requirements, such as user interface.

b) Accuracy: To determine if automated DTM extraction techniques do actually attain acceptable accuracies, the following tests have been devised for this purpose.

- Comparing DTM generated by softcopy system automated correlation methods, against DTM generated manually by a WisDOT operator using conventional plotters.
- Comparing DTM generated by softcopy system automated correlation methods, against DTM generated manually by a WisDOT operator using same softcopy system.

The WisDOT mapping accuracy standards require that error must not exceed 0.3 inch for 100% of direct observations, and same for at least 90% of indirect observations.

c) Speed: A substantial gain in time can be achieved if the accuracy attained by automated extraction methods are higher or at least comparable to that normally attained by an experienced photogrammetric operator.

Another criterion is defined by operator-system interaction time required for a full end-to-end stereophotogrammetric process. Interaction time is observed in the softcopy systems and compared with existing WisDOT standards for a typical mapping operation.

3. EVALUATION RESULTS

3.1 Data Management

The images used in this project were scanned in 22.5, 15, and 12.5 microns. The file sizes ranged 340+, 235+, and 100+ megabytes. This range of data volumes requires significant storage capabilities. There are some options to deal with such extensive amount of data for production environment. These include options such as central data bank, local area network, etc.

3.2 Eye Fatigue

The Intergraph and Helava systems utilize two different technologies for stereo viewing. In both cases, however, the WisDOT operators commissioned to examine the systems on site, i.e., UW Softcopy Photogrammetry Lab, expressed eye fatigue shortly after they used the systems. These operators were trained on, and routinely used, stereoscopic vision with optical components with continuous-tone hardcopy media. Optics of most conventional photogrammetric machines tend to be of high quality and craftsmanship, and hence very expensive. This is to ensure best visual quality, geometric accuracy, and above all, to enable the operator's eyes hours of focusing with minimal exhaustion. This can be generalized with operators of any conventional photogrammetric production facility. It is anticipated, however, that eye fatigue would be reduced as operators get used to the systems.

3.3 Accuracy

Comparing DTM generated manually by a DOT operator, using conventional plotter, against an automatically extracted DTM.

Model 1:

No. of DTM points Compared:	3046
Average Error Calculated:	0.7067 ft.
Standard Deviation:	1.063

Model 2:

Comparing DTM generated manually by a DOT operator, using Helava system, against two automatically extracted Comparing. Each was generated using different Ground Sampling Distance (GSD).

Manual DTM GSD =	25 ft.
Automatic DTM GSD =	25 ft.
No. of DTM points Compared:	744
Average Error Calculated:	-0.08 ft.
Standard Deviation:	1.68 ft.

Manual DTM GSD =	25 ft.
Automatic DTM GSD = then interpolated to 25ft.	12.5 ft.
No. of DTM points Compared:	744
Average Error Calculated:	-0.035 ft.
Standard Deviation:	1.28 ft.

There are two conditions under which the above accuracies were obtained. First, ground controls were minimal in all above models, and second, images were scanned at a 22.5 micron resolution. With photo scale at 1" = 300', one pixel is 0.25 ft on ground. Scanning with 12.5 micron resolution would have yielded 0.15 ft on ground, a significant difference.

3.4 Speed

An average of twenty percent of all models seems to be covered by elevated vegetation. Manual editing or manual extraction would be needed to obtain correct elevations of ground surface. This task is performed within the manual extraction/editing option available in most softcopy photogrammetric systems. The time needed for the additional manual extraction conforms to normal production standards of manual DTM extraction, assuming conventional photogrammetric systems. This is considered a loss in the speed and efficiency gained by the automated DTM extraction capability of the softcopy system.

Durations were estimated for operator-system interaction required for standard stereophotogrammetric procedures. An overall estimate was recorded for the entire end-to-end process in the Intergraph system. A detailed step-by-step breakdown periods were recorded for Helava. On both cases, interaction durations were superior to what is expected in a conventional photogrammetric machine.

3.2 Expertise

There is a shift in the kind of expertise required to operate a softcopy photogrammetric workstation. This expertise is a departure from what is normally required for conventional photogrammetric systems. Characteristics of this shift can be summarized as follows:

- For operators, less knowledge required in photogrammetric engineering, as a result of the "black box" concept, while more required in computers, Unix, windows, file systems, disk storage, stereo display, etc.
- For operators, basic system training is sufficient.
- For team leaders, high level training is required, to ensure smooth flow of operation when operators encounter a difficult, out-of-usual, situation with system software/hardware.

4. CONCLUDING REMARKS

-It has become obvious that softcopy photogrammetry technology does meet performance standards, even for large scale mapping, such as WisDOT photogrammetric operations.

-Substantial gain in speed can be achieved with softcopy photogrammetric systems, assuming other bottle necks, such as data volume, are handled efficiently.

-High end systems do provide at least comparable accuracy, when automated extraction methods work. However, substantial enhancements are needed to add some intelligence in the extraction process. Matching techniques tend to fail with simplest unexpected conditions.

-Conversion of conventional production facilities into softcopy operations requires significant re-training of operators. Cost issues, data volume, technical SW/HW support need also to be carefully considered.

ACKNOWLEDGMENT

This paper deals with aspects of a research project funded by the Wisconsin Department of Transportation, under project number 0092-35-57.

Évaluation de systèmes photogrammétriques à images vidéo : concept, méthode d'essai et résultats préliminaires

Mots-clés : Numérique, images vidéo, systèmes photogrammétriques, production, évaluation, essai

RÉSUMÉ

Le domaine de la photogrammétrie subit actuellement des mutations profondes attribuées principalement à la manière dont les données sont manipulées, analysées, et présentées. La mise au point de nombreux systèmes photogrammétriques à images vidéo est caractéristique de la direction de ces mutations. Un grand nombre des utilisateurs est encore très prudent devant ces nouveaux systèmes qui font leur apparition sur le marché et dont on vante les mérites. En outre, on peut dire sans se tromper que les entreprises d'exploitation de la photogrammétrie ont déjà investi des ressources considérables en machines de restitution, en équipement photographique et en formation des opérateurs. Par conséquent, il serait difficile de s'orienter dans une nouvelle direction qui non seulement rendrait le matériel actuel désuet, mais imposerait également de nouvelles immobilisations importantes. Il est évident qu'il n'existe pas encore de systèmes d'évaluation normalisés qui permettent de comparer certains systèmes à images vidéo à des postes de travail de photogrammétrie traditionnels. Une telle évaluation objective est nécessaire pour indiquer aux utilisateurs les avantages et les inconvénients des systèmes photogrammétriques à images vidéo. Le laboratoire de photogrammétrie à images vidéo de l'Université Wisconsin-Madison a entrepris un projet d'évaluation d'un certain nombre de systèmes photogrammétriques de ce type que l on trouve dans le commerce. La sélection de ces systèmes est déterminée par des facteurs précis, à savoir les contraintes budgétaires, les ressources restreintes en personnel qualifié pour effectuer les essais et la volonté de participation du fabricant. Le document présente les objectifs d'évaluation, la stratégie et les résultats préliminaires. En considérant certaines applications, les systèmes à images vidéo sont plus compétitifs, sinon supérieurs, aux systèmes photogrammétriques traditionnels. Toutefois, un certain nombre de facteurs précis doivent être soigneusement examinés comme, le volume des données, les compétences des opérateurs et l'automatisation.

STANDARDS FOR IMAGE SCANNERS USED IN DIGITAL PHOTOGRAMMETRY

Scott Miller and Alex Dam
Helava Associates Inc., 10965 Via Frontera, San Diego CA 92127

ABSTRACT

Digital photogrammetric products can be produced from virtually any type of imagery. Presently, most digital photogrammetric products are produced from scanned diapositives. The image scanner can typically be the most influencing factor on the accuracy of data collection. A Digital Photogrammetric Workstation (DPW) cannot recover from a poor quality scanner. This paper discusses methods for geometrically calibrating scanners and suggests that a standard should be made.

KEY WORDS: Scanner, Accuracy, Standards, Digital Image

INTRODUCTION

Photogrammetric products can be produced from virtually any type of image. The image can be from hand held 35 millimeter cameras, video cameras, satellite imagery, panoramic cameras and many others. The accuracy of products produced is largely affected by the input source. There is a general expectation that photogrammetric products produced from high quality mapping cameras, with distortions in the micron range, would produce products commensurate with input image quality. If we are to use scanners to digitize high quality films for photogrammetric data production, we must be careful to preserve radiometric and geometric quality if we are to compete with existing photogrammetric products such as analytical stereo plotters and high quality comparators. If one does not wish to preserve the fidelity of the input image and is cognizant of this (i.e., it is not required) then there is no problem. If one is not cognizant, and believes they are getting a product equivalent to the film camera output, then there is a serious problem. What standards are required? Should we have a rating system? In this paper, we are limiting our discussions to geometric quality, there is of course a very related issue of radiometric quality.

WHAT PHOTOGRAMMETRY EXPECTS

Today, many firms view photogrammetry in terms of C-factor or similar rules. Through experience and practice, firms learn what is achievable from a given instrument type.

C-factor = Flying Height of Image / Contour Interval of Map

Analytical plotters are typically achieving in the 1200 to well over 2000 range for C-factor. Stage accuracy's for first order analytical plotters are generally better than 3 microns. For the most part, a DPW should be limited in accuracy only by the quality of pixels given to it and the quality of control points used. Assuming the pixels are of good quality, algorithms and the human can easily measure to better than 0.5 pixels. Many studies prove that precision better than 0.25 pixels are routinely possible. To obtain accuracy, we need not just precision, but absolute knowledge of where those pixels are with-respect-to the original film focal plane of the camera. Much has been done in photogrammetry to assure good geometric fidelity of film when placed on the stage plate of an analog or analytical plotter. Why should photogrammetry require camera calibrations if the scanner does not require it? A scanner should preserve and adhere to the quality of analytical plotters and/or should state the accuracy to be expected under typical or standard operating conditions. If the scanners geometric quality is as good as the analytical plotter stages, then a user's C-factor concepts can be used for DPWs as well (assuming the pixel radiometric quality and size is good enough). If we scan an image with a 25 micron pixel size, and assume that measurements can be made to 0.25 of a pixel, we get a desired precision of 6.25 microns. This tells us that our scanner geometric error should be significantly less than 6 microns to achieve the best results and not unduly influence accuracy from a 25 micron pixel scan image.

General Scanner Types And Error Sources

Geometric qualities of scanned film primarily results from the scanner design of which there are several types. Stage types of scanners have a big impact on scan quality. Drum scanners that primarily are derived from the graphics art industries, typically use tension to hold the film flat against the drum. This combined with drum "roundness" make it difficult to maintain focus and geometric position in relation to the scan head. Flat bed stage designs can use hard optically flat cover plates to hold the film flat during scanning. This is an easier design to hold to higher geometric accuracy. Flat bed stages with cover plates are less susceptible to operator error in mounting the film. Scanners designed for a typical application are normally designed with sufficient accuracy for that application. In the graphic art applications, desk top flat bed and drum scanners are typical. Their accuracy is generally sufficient for the application they were designed for. The Sharp JX-600 uses a flat bed design but does not use a flat cover plate and does not use a high precision stage drive or low distortion lens. Its geometric accuracy is in the 80 micron range (Sarjakoski 1992). It produces reasonably good pixels and could be rated for photogrammetry, perhaps by stating its limitation on C-factor.

Some scanners offer calibration procedures using film or glass plate grids that can be measured. Some scanners state their positional accuracy, many do not. Some users will perform an affine transformation on the 4 or 8 fiducials of a diapositive to see how large the residuals are. This is sometimes misused to judge the accuracy of the scanning stage. Due to severe limitations on redundancy, and a lack of calibration points within the main scanning area of the film, this method should not be used to judge scanner accuracy.

The following is a partial list of error sources influencing the geometric accuracy of a scanned digital image:

X-Y stage or drum positioning: Is it repeatable? Can it be calibrated?, Does it change with temperature?

Film flatness: Does it stay flat? Is it optically flat? Can it bubble?

Radiometry: Are the pixels highly correlated? Is there a lot of noise? Is it out of focus? Is there enough dynamic range?

Lens distortion: Is it significant? Can it be calibrated?

Interior Orientation: Can it be performed? What is the measuring method?

Sensor Errors: Is the camera digital output linear? Can it be calibrated? Are the pixels square?

A manufacturer is not in a position to guarantee the accuracy of a scanned image. Too many operational factors come into play. The supplier of photogrammetric products is generally responsible for accuracy guarantees. The user of the scanner must be able to produce the accuracy statement. The manufacturer can supply the means by which the user can test the image accuracy.

A standard for stating the geometric accuracy for a scanned image, and a standard set of methods for achieving the statement of accuracy would be a big help.

Calibration Methods

Some scanner stages make routine use of calibration procedures to indicate and prove their geometric fidelity. These calibrations can be used at any time to analyze or set the geometric accuracy. Scanners that provide this knowledge, allow the user to calculate the scanners effect on C-factor or accuracy. The Helava DSW 100 and 200, The Zeiss-Intergraph PS-1, and others use this type of calibration to verify geometric quality. The Vexcell VX 3000 scanner accomplishes the same thing by providing grid calibration with each scanned image (Leberl 1992). Stable base films with calibrated grid lines or a glass grid plate are the most common methods of checking and calibrating X-Y stages. This is a fairly simple procedure, it gives confidence to the user and quality statements for the product.

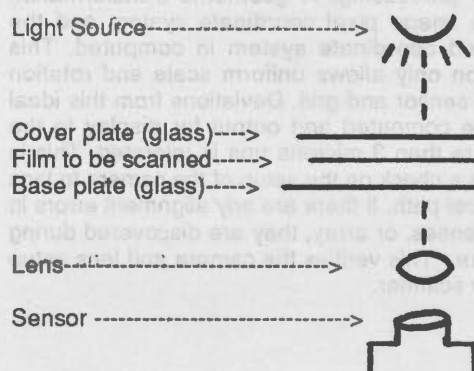
Besides the stage, the resulting pixel data needs to be checked for geometry. An accurate stage will not assure an accurate motion of the stage with-respect-to the sensor or vice-versa. The relationship of the sensor to the camera, and the camera to the stage needs checking. This can be short circuited by testing the resulting image. One method of doing this is by scanning a calibrated grid and then measuring the resulting image.

EXAMPLE CALIBRATIONS OF THE DSW 200

The DSW 200 is a photogrammetric film image scanner produced by Helava Associates Inc. Like the DSW 100, this scanner uses a 10 X 10 inch flat bed stage with 1 micron encoders. It typically calibrates about 2 microns Root Mean Square (RMS) error in X,Y position. It Scans an image using a 2k X 2k area array sensor at physical pixel sizes between 7 and 15 microns. It outputs pixels to the host computer at approximately 1 million pixels per second. The pixels for each area (about 2kX2k) are tiled together by the host computer to generate a seamless image of up to 10 by 10 inches.

Scanning Mechanics

The X-Y stage is commanded to move to a given position above the sensor and comes to rest before an image is acquired. A highly uniform light source is transmitted through the cover plate, through the film, through the base plate, through an imaging lens, and to the sensor elements in the camera as shown in the figure below.



By calibrating the relationship of the sensor array to the stage coordinate system, we can compute each position of the stage for collecting an entire array of 2k X 2k images to cover the entire stage area. Each stage position is accurate to less than 3 microns RMS, and each pixel position within the sensor is accurate to less than 3 microns RMS. How do we know this?

Stage Calibration

The X-Y stage positioning of the scanner is calibrated using an etched glass grid plate. It is on this grid as well as the precision and high stability of the opto-mechanics that stage accuracy is based. The grid lines etched on the grid plate are accurate to less than 1 micron as guaranteed by the manufacturer. This grid plate has a grid spacing of 20 mm in X and Y. Based on this calibrated data, DSW 200 will automatically find every grid intersection, measure it, and record the physical stage coordinates to the nearest encoder micron. Image processing algorithms are used to find and measure the intersection to stage precision.

A table lookup of stage corrections can now be built between the physical (encoder) stage coordinates and the calibrated grid plate coordinates. This lookup table of corrections is now used to produce calibrated stage coordinates for any commanded stage position. This calibration can be performed at any time and takes about 30 minutes. Additionally, there is a "stage verification" mode which measures a subset of the grid intersections and tells the user if there is any problem with the current calibration. This procedure takes about 5 minutes. The verification compares the calibrated stage coordinates against the actual grid plate coordinates and shows the user the RMS error in X and Y. If this exceeds their desired threshold, the stage is recalibrated. The stage verification procedure is recommended about once a month and recalibration is rarely necessary. Due to a well proven design, the stage holds calibration for very long periods of time.

Sensor Geometric Calibration

The sensor image of the DSW 200 is also calibrated for geometric accuracy to the same level as the stage. This can be done by another glass grid plate that has a grid line spacing of 5 mm and whose intersections are also accurate to 1 micron. A 6 by 6 grid of intersections will then appear on the focal plane of the 2k by 2k area array when using an optical pixel size of 12.5 microns. This grid data again permits us to automatically find and measure each grid intersection to sub-pixel accuracy using image processing. A geometric transformation between the image pixel coordinate system and the calibrated grid coordinate system in computed. This transformation only allows uniform scale and rotation between the sensor and grid. Deviations from this ideal geometry are computed and output for display to the user. No more than 3 microns rms is tolerated. This is performed as a check on the setup of the camera to lens to stage optical path. If there are any alignment errors in the stages, lenses, or array, they are discovered during this procedure. This verifies the camera and lens setup for each new scanner.

Checking The Final Image

The scanned image is an array of pixels residing in the computer or on the hard disk. The image can be displayed and measured. If we assume a maximum scan of interest to cover about 220mm x 220 mm (9 inch x 9 inch), then we should have a reference target of that size to scan. A calibrated grid plate or film transparency seems ideal, with intersections at a spacing sufficient to uncover local distortion but not so frequent as to be impractical to measure. The technology to produce such accurate grids is readily available so no new wheel needs inventing. This proposed reference would be scanned and an image file produced. Next in this case, every grid intersection would be measured by the human placing a cursor on each perceived grid intersection or by using image processing. When using the human and a computer monitor, we would typically zoom the image by a factor of 2 or 4 to help place the cursor to sub-pixel positions on the grid intersections. However, an algorithmic processing method would be preferred due to its ease, speed and repeatability if one could be agreed upon. If universally agreed upon algorithms were employed, then a real scanner standard would begin to be possible.

Let us assume the grid has intersections every 20 millimeters. That would require measuring about 100 grid intersections over the entire image. With these measurements, a 4 parameter transformation from pixels to grid coordinates could be computed. This transformation physically models rotation, translation, and a scale change between the grid coordinate system and the pixel coordinate system. The RMS of the residuals from this transformation is a good indicator of accuracy of the scanned image because the redundancy is so high and testing was done throughout the scanned image. This procedure can be performed by us the manufacturer, or by the user. This permits a statistically rigorous check on the resulting image geometry directly.

RECOMMENDATIONS

Since the user of the image scanner or scanned image data is really in control of the geometry of the image, and since it is a relatively easy procedure to verify the quality of a scanned image, we would recommend that the user of the scanner test and produce a standard accuracy statement. We would also recommend that the manufacturer of the scanner provide the software tools necessary to measure a calibration plate or film so the user can produce the accuracy statement. We would recommend that the professional associations establish the specific methods by which this should be done and the minimum content of the statement. The manufacture or the user could supply the calibration film or plate. A large 9 by 9 inch scan should be checked with a minimum of 20 points spread throughout the scan area. This would permit sufficient redundancy when checking the quality using a 4 parameter transformation. The resulting accuracy statement should minimally state the RMS in X and Y directions along with the maximum fit residual.

References

Leberl, Franz W., 1992, Precision Scanning Of Aerial Photogaphy: ASPRS Technical Papers, 1992 ASPRS-ACSM Annual Convention, Vol.1 pp. 247-252.

Mikhail, Dr. Edward M., 1992, Quality Of Photogrammetric Products From Digitized Frame Photography: 1992 ISPRS, International Archives Of Photogrammetry And Remote Sensing, Vol. XXIX, Part B2, Commission II, pp. 390-396.

Sarjakoski, Tapani, 1992, Suitability Of The Sharp JX-600 Desktop Scanner for The Digitization of Aerial Color Photographs: 1992 ISPRS, International Archives Of Photogrammetry And Remote Sensing, Vol. XXIX, Part B2, Commission II, pp. 79-86.

«Normes visant les balayeurs d'image utilisés en photogrammétrie numérique»

RÉSUMÉ

Presque tous les types d'images peuvent servir à la réalisation de produits de photogrammétrie numérique. À l'heure actuelle, la plupart de ces produits sont obtenus par balayage de diapositives. Le balayeur d'image peut normalement constituer le facteur le plus déterminant de la précision de la collecte de données. Un poste de travail de photogrammétrie numérique ne peut récupérer d'un balayeur de mauvaise qualité. Cet article traite de méthodes pour l'étalonnage géométrique des balayeurs et propose qu'une norme soit établie.

Mots-clés : Balayeur, précision, normes, image numérique

PSEUDO-STEREO DIGITAL PHOTOGRAMMETRY

Eugene E. Derenyi and Ying Chen*

Department of Geodesy and Geomatics Engineering, University of New Brunswick
P.O. Box 4400, Fredericton, N.B. Canada E3B 5A3

* Now at The Department of Geomatics Engineering, University of Calgary
Calgary, Alberta, Canada T2N 1N4

Commission II, Intercommission Working Group II/III

KEY WORDS: Digital photogrammetry, Image matching, GIS, Digital mapping.

ABSTRACT

A so called pseudo-stereo technique was developed for soft-copy mapping which allows spatial positioning without the need for stereo viewing or a digital elevation model. Digitization is performed monoscopically on the display screen of a computer graphics workstation in one of the images of a stereo pair. Corresponding points in the second image are located by area based least squares image matching. The object space coordinates are then computed by space intersection. A close initial approximation of the matching position is assured by enforcing the epipolar geometry as a constraint and by an a priori estimation of the elevation of the feature to be mapped. This mapping technique was successfully implemented and tested in a commercial geographic information system.

1. INTRODUCTION

Soft-copy, digital photogrammetry is revolutionizing map compilation. The production of general purpose basic maps, such as the federal, provincial or municipal map series, will probably remain in the hands of professional mapping organizations in the foreseeable future. Natural resource inventory maps, thematic maps, special purpose maps, and alike can now be produced by professionals and technologists in various disciplines, who are also the users of these products. With the help of mapping software, these individuals can extract, in a computer graphics workstation, the information they need from digital images. Extensive knowledge in photogrammetry and cartography is no longer necessary. This desk top mapping is analogous to desk top publishing which is now widely applied.

The first commercial versions of all-digital photogrammetric restitution instruments appeared in the 1980s, and several models were exhibited at the 1992 ISPRS Congress. Currently, the high performance workstations are rather expensive, and are affordable only by large organizations. There are also some technical problems to be overcome before soft copy images replace the diapositive. More powerful processors, larger and faster memories, higher resolution display screens, and the development of a CCD camera with a performance comparable to a cartographic aerial film camera are needed. Electronic technology is advancing at a rapid pace, thus it will not take all-digital photogrammetric systems thirty years to come to age, as was the case for the analytical plotters. In the meantime, attention should be focused on the so called secondary mapping operations like, resource inventory, thematic, customized mapping and map updating. Existing technology can serve these applications well, in a cost effective manner.

Secondary mapping is performed under different conditions and requirements than primary basic mapping, in the following respects:

- Planimetric information extraction is the main goal.

- Map accuracy standards are less stringent than in primary mapping.
- Secondary mapping implies that a primary digital base map already exists which can provide a dense network of reference features for the restitution of images.
- Resolution of digital images used in most secondary mapping applications need not be as high as for primary mapping.

The requirements which a secondary soft copy mapping facility must fulfill can be summarized as follows:

- All operational procedures must be easily understood and able to be followed by users with limited photogrammetric or cartographic skills
- Image restitution must be based on sufficiently rigorous mathematical models to assure that the map accuracy standards and specification are met.
- Means must be provided for the easy integration of the mapping output with existing data bases.

The pseudo-stereo mapping technique is designed to satisfy these requirements.

2. PSEUDO-STEREO IMAGE RESTITUTION

2.1 The Technique

Monoscopic measurement and digitization in images displayed on the screen of a cartographic workstation form the basis of this mapping technique. A significant element of this scheme is its implementation in a geographic information system (GIS). Thus, information stored in the GIS can serve as ancillary data to support the mapping operation while the newly acquired information can readily be merged with existing files.

Image monoplotting is the simplest mapping scheme which can be offered to non photogrammetrists. In the past, single image photogrammetry has been used routinely for transferring details from photographs onto maps in instruments built on the camera lucida principle. The residual effects of tilt, relief displacement and other geometric anomalies inherent in photographs are, however, limiting factors in the application of these devices. Possible means to alleviate these problems in soft copy photogrammetry are to perform the digitization in a digital orthoimage or to correct the relief displacement at the digitized points with the help of a digital elevation model (DEM). Pseudo-stereo mapping is design for the case when orthoimages or DEMs are not available, and, although the digital images provide stereoscopic coverage, the workstation is not equipped for stereo viewing.

2.2 The Procedure

The pseudo-stereo mapping technique has been implemented in a GIS [Chen, 1993] which is capable of handling digital images in raster form as well as digital maps in vector form [Derenyi, 1991]. The main steps are as follows (Figure 1):

1. The exterior orientation parameters of both photographs of a stereo pair are determined by space resection. Image coordinates of the ground control points (GCPs), needed for this purpose, can be measured monoscopically on the display screen. When a digital map of the area is available in the GIS, the GCPs can be picked in a double window display of the image and the map. An alternate step is to acquire the orientation parameters from some other sources such as aerotriangulation.
2. The planimetric features to be mapped are digitized in Photo 1 on the screen. This is achieved by free hand cursor control using a mouse. Zooming, planning and real time image enhancement facilities assist in this operations.
3. Corresponding image points are located in Photo 2 by image matching. This operation is performed in batch mode after a group of points has been digitized.
4. The object space coordinates of the digitized points are computed by space intersection and stored in a new map file or added to an existing file. This step is also a batch operation.

2.3 Image Matching

The key issue in pseudo-stereo mapping is to locate the corresponding point in the second image of a stereo-pair by image matching. There are several techniques available to perform this task, which fall into two categories: area-based and feature-based matching. The area-based least squares image matching (LSM) technique has been selected for this development. LSM derives a shaping function from one image window to another by minimizing the gray level value differences between the two matching windows. Briefly, this is done by forming an observation equation for each pixel being matched at a particular position of the target array as it moves in the search area. Each observation equation compares the gray level values of the pixels being matched to estimate geometric and radiometric differences between the two images [Rosenholm, 1986].

LSM has several advantages:

- High precision can be obtained. Precision of up to 0.05 pixel has been reported in the literature.
- The geometric and radiometric differences between two images can be modeled.
- The solution is highly redundant, even for small window size, which provides a strong least squares solution.
- The information obtained at one matching position can be used as a guide for the next iteration, which avoids a systematic search. LSM is therefore computationally efficient and is suitable for a real time solution.

LSM also posses certain problems:

- The algorithm, which has been developed by the use of Taylor's expansion to linearize the observation equations, excludes the non-linear higher order terms from the solution process. Therefore, a good initial approximation is needed for the matching position .
- Multi-solution (non-unique solution) may occur in areas where repetitive patterns and high frequency texture exist.

In the development of the pseudo-stereo mapping technique these problems have been largely overcome by predicting the close vicinity of the matching position. This is accomplished by applying the epipolar geometry as a constraint, and introducing a close approximation for the elevation of the feature point to be mapped. In the absence of other information, the average terrain elevation, calculated from the control points, is accepted as the initial value. A contour map and familiarity with the general trend of the topography could facilitate the making of a better selection. This elevation is later refined as the correlation and the determination of the object space coordinates progresses and more exact elevation values are generated [Chen, 1993]. The procedure followed in the computation of the object space coordinates of the digitized points is shown in Figure 2.

3. PERFORMANCE TEST

3.1 Test Materials and Method

The pseudo-stereo mapping technique was tested on a stereo pair of black-and-white photographs, covering the City of Fredericton and vicinity. The photographs were taken at a scale of 1:35 000, with a wide angle camera. The downtown area, which spreads along the south shore of the St. John River, is essentially flat at an elevation near sea level. From there the terrain has a steady incline and reaches an elevation of 130 m at the city limit.

Digital images were generated by scanning the paper prints in a Hewlett-Packard ScanJet Plus document scanner. The contrast and brightness parameters of the scanner were adjusted before scanning to equalize the mean gray level values in the two images. Considering the limited disk space available in the computer, the scanning resolution was restricted to 300 dots-per-inch (dps) or 118 dots-per-cm, which resulted in a pixel size of 85 μm in the image and 3.0 m on the ground. The radiometric resolution was set to 256 gray levels.

GCPs for the determination of the exterior orientation elements were obtained from 1:1 200 scale maps. The contour interval is 1.0 m overall, but in areas where the elevation is below 15 m, the interval is 0.5 m.

A total of 120 feature points of different types were digitized on the screen in Photo 1. The features selected were residential streets, buildings, a bridge, railway line and a sports field. The ground coordinates of each were then computed using the correlation and the space intersection software, and compared with the coordinates measured on the 1:1 200 scale maps [Chen, 1993].

3.2 Test Results and Discussion

The results of the test are summarized in Table 1.

The RMSEs are the root mean square errors of the planimetric coordinates obtained by the pseudo-stereo mapping technique with respect to those measured in the 1:1200 scale maps. The notation "% < 1.5 m", etc. means the percentage of the check points with differences smaller than 1.5 m, etc. The photographs for this test were taken from a block flown for 1:10 000 scale mapping. Therefore the assessment of the results are based on the degree of compliance with the map accuracy standards applicable at that scale.

The accuracy specification for the "Urban and Resource Digital Map Base" in the "Land and Water Information Standards Manual" states [NBGIC, 1991]:

Ninety percent (90%) of all "well defined features" must fall within the positional accuracy ...[2.5m]. Well-defined features are those whose positional accuracy is not adversely affected by vegetative cover. Accuracy of the digital data (point, line, area) can be defined as the difference between the position of the associated data in the digital file and the real position of the represented features on the earth.

Although this requirement is rather strict, 87% of the well defined feature points satisfied it. This is only 3% below the 90% tolerance. For graphical map production, the generally accepted tolerance in North America is 0.5 mm at publication scale, at the 90% confidence level. This corresponds to 5.0 m for the 1:10 000 scale. All points were within this tolerance.

The following points should be considered to put the results in proper perspective:

1. The pixel size of the digitization was 3.0 m on the ground, which is larger than the 2.5m tolerance set by the map accuracy standards.
2. The stage plate of the scanner used was 8.5 x 14 inches, which is narrower than the size of the photographs. Thus, two of the primary corner fiducial marks and one of the secondary marks along the edges of the photographs were missed. This introduced an uncertainty in the interior orientation.
3. The particular scanner employed is only suitable for digitizing opaque photographic prints. The poor dimensional stability of paper is of course a source of significant distortion.
4. Because of storage limitations in the computer, only the overlapping areas of the photographs were kept on the files and used for selecting GCPs. Therefore, the geometry of the bundle of rays in the space resection were not the most favorable.
5. The 1:1 200 scale maps used for comparison was accepted as error free.

Under the above circumstances the test results can be considered satisfactory and accepted as a proof of the feasibility of the pseudo-stereo mapping technique developed. Further tests are of course needed to fully validate and improve this technique.

4. CONCLUSIONS

The pseudo-stereo mapping technique is applicable when no orthoimages or DEMs are available, or when the DEM is out of date and the relief displacement is too significant to be ignored. This technique has been successfully implemented in a GIS as proven by the outcome of an experiment. The ability to perform soft copy mapping in a GIS has the advantage that both the image and the map reside and are manipulated in the same workstation.

It is true that the results of the test do not entirely satisfy the map accuracy standards set for the digital base map production. Firstly, the experiment was conducted under less than ideal conditions for precision mapping, as explained earlier. Secondly, the objective of this development was to provide tools for thematic mapping, performed by non-photogrammetrists, when accuracy requirements are not as stringent as in basic mapping. Thirdly, simplicity and low-cost were also objectives. This explains the reason for omitting the need for stereo vision and for the use of an inexpensive document scanner and paper prints. Corrections could have been included to reduce the various distortions encountered but this would complicate the procedure and may prove to be beyond the comprehension of non-photogrammetrists. The most feasible way of improving the accuracy would be to digitize the diapositives in a larger, more precise scanner, at a higher density.

REFERENCES

Chen, Y, 1993. *Development of Pseudo-stereo Method for Digital Map Revision.* M.Sc.E. thesis, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B., Canada, 97 pp.

Derenyi, E., 1992. Design and development of a heterogeneous GIS. *CISM Journal ACSGC*, 45(4):561-567.

NBGIC (1991). "Province of New Brunswick Land and Water Information Standards." New Brunswick Geographic Information Corporation, Fredericton, N.B., Canada.

Rosenholm, D. (1986). *Accuracy Improvements in Digital Matching*. Phot. Reports, No.2: 52, Dept. of Photogrammetry, the Royal Institute of Technology, Stockholm.

Table 1: Results of the pseudo-stereo mapping experiment.

Feature	No. points	Mean (m)	RMSE(m)	% < 1.5m	% < 2.5m	% < 5.0m
Residential streets	62	1.41	1.60	58.1	88.7	100
Buildings	32	1.63	1.85	50.0	87.5	100
Bridge	6	1.15	1.29	83.3	100	100
Railway	9	1.18	1.35	77.8	88.9	100
Sport field	11	1.47	1.72	72.7	72.7	100
All	120	1.44	1.65	60.0	87.5	100

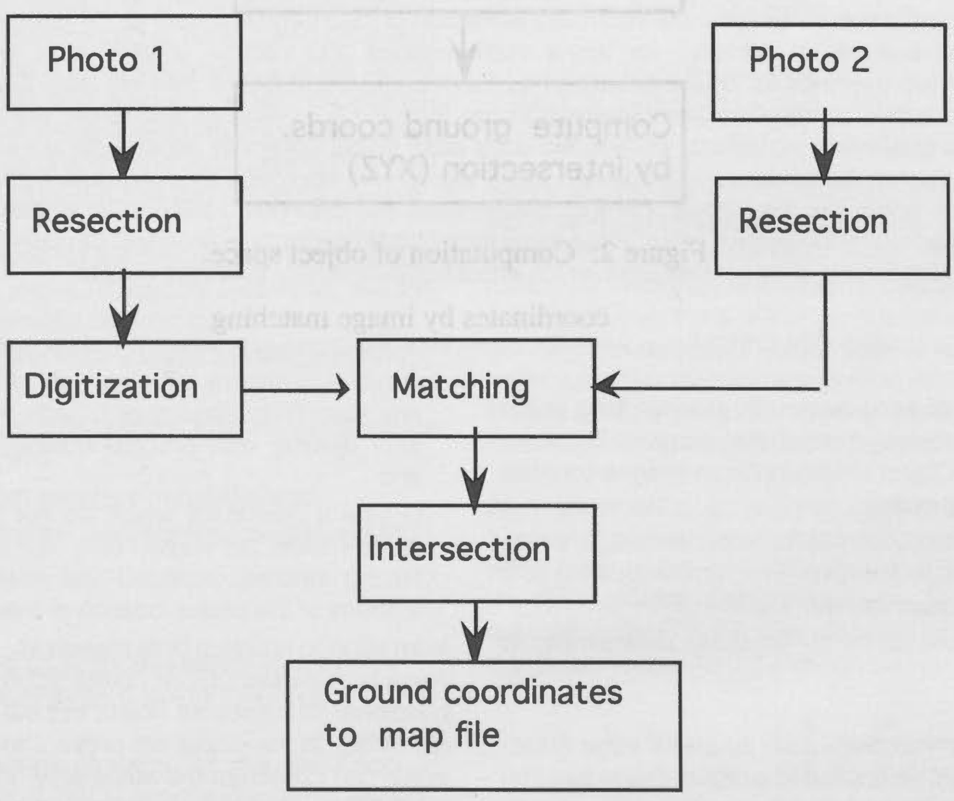


Figure 1: Pseudo-stereo mapping procedure.

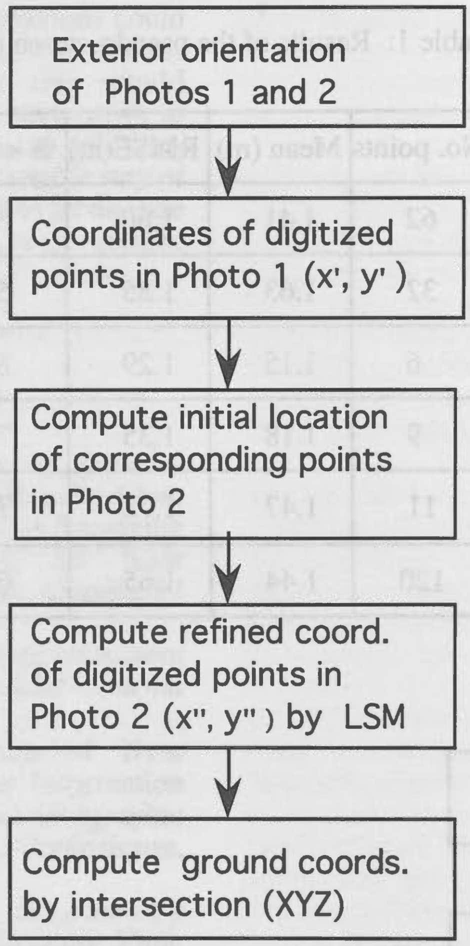


Figure 2: Computation of object space coordinates by image matching

Photogrammétrie numérique pseudo-stéréoscopique

RÉSUMÉ

Une nouvelle technique de cartographie dite pseudo-stéréoscopique, où les images n'apparaissent qu'à l'écran ou sur support fugitif, permet maintenant d'effectuer du positionnement spatial sans recourir à la visualisation stéréoscopique ou à un modèle numérique d'altitude. Grâce à ce système, la numérisation s'effectue sur l'écran d'un poste d'infographie en mode monoscopique, c'est-à-dire à partir d'une seule des images d'un couple stéréoscopique. Les points correspondants de l'autre image de la paire sont repérés grâce à la méthode d'appariement d'images des moindres carrés. Les coordonnées de l'objet spatial sont ensuite calculées par la méthode d'intersection spatiale. On peut obtenir une estimation initiale relativement fiable de la position d'appariement par l'application des règles de la géométrie épipolaire, et par une estimation a priori de l'altitude de la caractéristique à cartographier. Cette technique de cartographie est maintenant reconnue, puisqu'on l'a intégrée à un système commercial d'information géographique qui a depuis fait ses preuves.

SEMI-AUTOMATIC MONOPLOTTING ON A DIGITAL PHOTOGRAMMETRIC STATION

Peggy Agouris, Dirk Stallmann, Haihong Li

Institute of Geodesy & Photogrammetry
Swiss Federal Institute of Technology
ETH - Hoenggerberg
CH-8093 Zurich, Switzerland
Ph.: +41-1-6333054, Fax: +41-1-3720438
e-mail: peggy *or* dirk *or* li @p.igp.ethz.ch

ISPRS Commission II, Intercommission II/III

KEY WORDS: Digital Photogrammetry, GIS, Automation, Monoplotting, Object Extraction, Matching

ABSTRACT

This paper deals with the subject of semi-automatic object extraction from digital imagery in monoscopic mode. Addressed topics include conceptual, algorithmic and implementational issues, performed experiments, as well as discussion of encountered problems and adopted solutions. In particular, we present the algorithm and obtained experimental results for the semi-automatic extraction of road networks from SPOT imagery using wavelet-transformed images. In addition, for larger scale images and various object types, we present the comparison of two methods for object extraction we have applied and tested, namely least squares template matching, and active contour models (Snakes). In the former case, edge locations are precisely identified based on local gray value variations, while in the latter case global continuity constraints are enforced to produce meaningful results. Finally, we propose a novel algorithmic approach for semi-automatic object outline detection which combines the strong points of the previous two methods. Least squares matching provides again the mathematical foundation while at the same time global continuity is enforced through the introduction of object-type-dependent shape constraints which ensure geometrically coherent results.

1. INTRODUCTION

The objective of semi-automatic monoplotting is the extraction of objects from digital imagery in monoscopic mode. Object extraction from images consists of the following phases:

- ❑ *identification* of an object within an image, which involves image interpretation, understanding and object classification, and
- ❑ *tracking* the object by precisely determining its outline.

It is well known that there exist no *global* edge detectors which could be applied to a digital image function to both identify and track edges with sufficient success. Instead, one can witness a trade-off between *reliability*, which expresses a measure of the qualitative accuracy associated with identification, and *precision* which expresses a measure of the geometric accuracy associated with tracking. According to these measures, one can classify existing operators into two broad categories [Fischler et al., 1981]:

- ❑ *type I operators*, offering high reliability in properly identifying classes of objects without particularly dealing with precise outline determination, and
- ❑ *type II operators*, which do not aim at reliable identification, but instead offer high precision in detecting outlines, provided that adequate approximations of the object location are available.

In an effort to optimize both measures, operators from these two classes can be combined in complex computational strategies for object extraction [Suetens et al., 1992]. In this paper we present several methods which can constitute the automated object extraction core within a broader digital photogrammetric semi-automatic monoplotting strategy.

2. SEMI-AUTOMATIC MONOPLOTTING

In semi-automatic monoplotting, the identification task of a type I operator is performed manually on a single image, while a special automated digital module performs the tracking task of a type II operator. More specifically, a human operator is used to identify an object from an on-screen display of a digital image, select the particular class to which the current object belongs (e.g. road, house etc.) and provide a rough approximation of the object outline. Typically, this approximation consists of loosely identifying on-screen nodes (e.g. corners for houses, breakpoints for curvilinear objects etc.) of the outline. Subsequently, these pieces of information are used as the necessary approximations for the automatic, precise edge positioning task. By repeating this process, any objects within an image can be identified and precisely positioned. The degree of automation varies according to the extent of the required human operator contribution (e.g. how many nodes have to be provided for successful object extraction and how close to the actual outline breakpoints).

Judging from experience in both analytical photogrammetric data collection and digital image feature extraction, such use of a human operator within the broader object extraction strategy is considered optimal. Humans perform the identification task flawlessly and almost effortlessly, and thus, their intervention optimizes achieved accuracies without imposing time burden. At the same time, the task of precise object outline positioning and tracking, which experience shows to be the most time-consuming and error-prone part of photogrammetric data collection, is performed automatically in a fast and objective manner.

In the next sections we will present the mathematical foundation and implementational issues for the following semi-automatic object exraction methods:

- ❑ road extraction from wavelet transformed SPOT imagery,
- ❑ the use of active contour models (snakes) for outline extraction,
- ❑ least squares template matching, and
- ❑ globally enforced least squares template matching.

3. ROAD EXTRACTION USING WAVELET TRANSFORMED SPOT IMAGES

The semi-automatic strategy for the extraction of road networks from SPOT images combines a wavelet decomposition for road sharpening with a linear feature extraction algorithm based on dynamic programming (Fig. 1).

Wallis filter preprocessing is used to enhance the available image and facilitate subsequent object

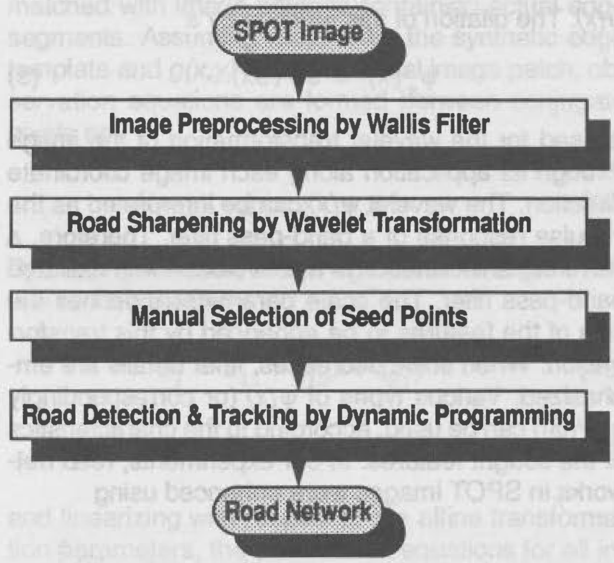


Fig. 1: Semi-automatic road extraction strategy

extraction processes by locally forcing the gray value mean and contrast (dynamic range) to fit certain target values. Filtering is performed in image blocks of $k \times l$ pixels by modifying the original gray value $g(i,j)$ of a pixel to a new gray value $f(i,j)$, as

$$f(i,j) = \frac{[g(i,j) \quad m_g]\, cs_f}{cs_g + (1-c)\, s_f} + bm_f + (1-b)\, m_g \quad (1)$$

where m_g, m_f are the old (original) and new (target) block mean gray values respectively and s_g, s_f are the old and new block gray value standard deviations. The histogram enhancement parameters c and b, which respectively affect contrast and brightness, are selected such that image noise is limited. A gray value transformation for any pixel is performed by bilinearly interpolating the filtering parameters of its four neighboring image blocks [Baltsavias, 1991].

After preprocessing, *wavelet transformation* is applied on the image to emphasize features of interest, while suppressing other details [Grossmann & Morlet, 1984], [Mallat, 1989]. A particular wavelet $\Psi(\omega)$ is uniquely defined in the frequency domain by specifying its associated filter function $H(\omega)$ as

$$\Psi(\omega) = G\left(\frac{\omega}{2}\right)\Phi\left(\frac{\omega}{2}\right) \quad (2)$$

where

$$\Phi(\omega) = \prod_{p=1}^{\infty} H(2^{-p}\omega) \quad (3)$$

$$G(\omega) = e^{-i\omega}\,\overline{H(\omega+\pi)} \quad (4)$$

In spatial domain, the corresponding wavelet will be $\psi(x)$. The dilation of this function by s

$$\psi_s(x) = s\psi(sx) \quad (5)$$

is used for the wavelet transformation of the image through its application along each image coordinate direction. The wavelet $\psi(x)$ can be interpreted as the impulse response of a band-pass filter. Therefore, a wavelet transformation is a convolution with a dilated band-pass filter. The scale parameter s defines the size of the features to be enhanced by this transformation. When scale decreases, finer details are emphasized. Various types of $\psi(x)$ (or correspondingly of $\Psi(\omega)$) can be used, according to the characteristics of the sought features. In our experiments, road networks in SPOT images were enhanced using

$$H(\omega) = \left(\cos\frac{\omega}{2}\right)^3 \text{ and } \Psi(\omega) = -\omega^2\left(\frac{\sin\frac{\omega}{4}}{\frac{\omega}{4}}\right)^5 \quad (6)$$

The wavelet transformation of an image generates a new image version, in which road networks are more prominent. These networks are then precisely identified through *dynamic programming*, which is a general multistage optimization technique to solve problems by maximizing a merit function (or minimizing a cost function) [Ballard & Brown, 1982]. Its application to edge extraction involves the definition of a function which embodies the notion of "best edge", and the solution of an optimization problem to extract the actual edge. A such function can be defined as

$$h(x_1, ..., x_n) = \sum_{k=1}^{n} s(x_k) + \alpha \sum_{k=1}^{n-1} q(x_k, x_{k+1}) \quad (7)$$

where $s(x_k)$ expresses edge strength, $q(x_k, x_{k+1})$ expresses change in edge direction between two successive edge pixels, and α is a negative constant. By evaluating the above function for possible edge paths, we select as edge the sequence of eight-connected pixels which maximizes the merit function. The disadvantages of this approach can be synopsized as follows:

- ❑ time consuming due to pixel-by-pixel tracking,
- ❑ sensitive to noise,
- ❑ difficulty in bridging edge gaps, and
- ❑ appropriate only for low curvature edges.

To overcome the above disadvantages, an edge can be extracted as a set of n smaller segments defined by the seed points $P^o_0, P^o_1, ..., P^o_n$ which are manually provided on-screen. The merit function is then expressed as the summation of smaller terms

$$h(P_0, P_1, ..., P_n) = h_0(P_0, P_1) + h_1(P_1, P_2) + \\ + ... + h_{n-1}(P_{n-1}, P_n) \quad (8)$$

The original seed positions can be altered during the solution. A preset value δ defines the maximum allowable change in the position of seed points $(|P_i - P^o_i| = |V_i| \leq \delta)$ and it is used to eliminate tracking errors associated with the selection of misleading local maxima. Each term $h_i(P_i, P_{i+1})$ expresses how well a path in the image connecting points P_i and P_{i+1} satisfies the merit function.

Applied to road extraction, the definition of the associated merit function is based on four observations:

- ❑ a road pixel is lighter than its neighbors on both road sides,
- ❑ a road is usually smooth and of limited curvature,
- ❑ gray values along a road usually do not change very much within a short distance, and
- ❑ road width does not change significantly.

These four conditions are formulated and combined as a merit function

$$h(P_i, P_{i+1}) = \sum_{P_k \in C_i} [S(P_k) + \alpha D(P_k) + \beta T(P_k)] \\ + \kappa Q(P_i, P_{i+1}) \quad (9)$$

where $S(P_k)$ denotes road strength, $D(P_k)$ the local variance of strength, $T(P_k)$ the local texture, $Q(P_i, P_{i+1})$ the difference of directions, and α, β, κ are constants.

This multistage optimization problem is solved by dynamic programming. The advantages of the modified optimization method (*improved dynamic programming*) are:

- ❑ efficient handling of long curves,
- ❑ robustness in the presence of noise, and
- ❑ ability to bridge edge gaps.

4. ACTIVE CONTOUR MODELS

Active contour models, also known as *"snakes"*, are energy minimizing splines, guided by shape and radiometry forces to fit to, and thus identify, edges in digital images [Kass et al., 1988]. Their implementation is semi-automatic, with an operator manually providing the necessary initial edge approximations in the form of a few seed points. The desired edge behavior is expressed in energy terms and assuming that local energy minima correspond to object boundaries, edges are identified by minimizing an energy function E. This function comprises of two terms, one radiometric (E_p), and another geometric (E_g):

$$E = E_p + \lambda E_g \quad (10)$$

Parameter λ expresses a measure of the roughness of the initial edge estimate. The radiometric energy part assumes that the first derivative of image intensity in the direction of the gradient is extremal at step edge points. The geometric part defines the form of the edge contour to be a cubic spline function, continuous in first and second derivative and thus enforcing its smoothness. The total energy is minimized through an optimization procedure which forces the snake to approach the actual edge contour (Fig. 2). Iterative convergence is achieved by mathematically simulating the behavior of a deformable body embedded in a viscous medium and solving the corresponding dynamics equation [Fua & Leclerc, 1990].

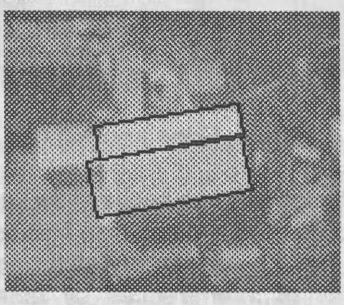


Fig. 2: Edge contour detected by snakes

Active contour models present several advantages:

- They are able to bridge radiometric gaps and weak regions because they are using global information.
- They can be applied for the extraction of both open and closed edge contours.
- They can successfully follow the contour around corners by relaxing the second derivative continuity constraint of their geometric energy part.
- Their mathematical foundation can be extended and customized by the use of additional geometric constraints to fit specific and various object types (e.g. parallel curves for road extraction).
- Snakes can be extended towards 3-D object extraction by using an underlying DTM or by including camera models.

5. LEAST SQUARES TEMPLATE MATCHING

This modified matching method, used for edge detection and tracking, is based on *least squares matching*. A synthetic edge pattern is introduced as the reference template which is to be subsequenlty matched with image patches containing actual edge segments. Assuming $f(x,y)$ to be the synthetic edge template and $g(x,y)$ to be the actual image patch, observation equations are formed between conjugate pixels as

$$f(x, y) - g(x, y) = e(x, y) \quad (11)$$

By relating template and image patch through an affine transformation

$$x_i = a_{11} + a_{12}x + a_{21}y \quad (12)$$

$$y_i = b_{11} + b_{12}x + b_{21}y \quad (13)$$

and linearizing with respect to the affine transformation parameters, the observation equations for all involved pixels can be written in matrix form as

$$-e = Ax - l \quad (14)$$

where l is the observation vector containing gray value differences of conjugate pixels, x is the vector of unknowns consisting of the affine transformation parameters, and A is the associated design matrix including the derivatives of the observation equations with respect to the unknowns [Gruen & Baltsavias, 1988]. The least squares matching solution is then obtained by minimizing the squared sum of gray value differences

$$\hat{x} = (A^T PA)^{-1} A^T Pl \quad (15)$$

and a new position of the image window is determined as the conjugate of the template through the updated affine transformation parameters (Fig. 3).

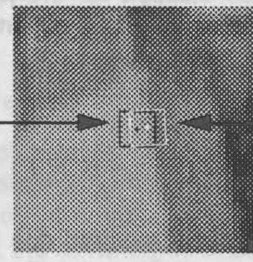


Fig. 3: Visualization of least squares template matching

However, due to the particular gray value distribution of the edge patches, a full set of affine transformation parameters cannot be obtained [Gruen & Stallmann, 1992]. Instead, only two shift and one rotation parameter relating template and image patch can be determined through the least squares solution. Therefore, either the affine transformation describing the relationship between template and image patch has to be substituted by a simpler three-parameter one

$$x_i = \Delta x + (x\cos\varphi - y\sin\varphi) \quad (16)$$

$$y_i = \Delta y + (x\sin\varphi - y\cos\varphi) \quad (17)$$

or, equivalently, three shape constraints should be introduced in the solution to effectively reduce the number of affine transformation parameters. These three constraints are formulated as

$$a_{12} = b_{21} \quad a_{21} = -b_{12} \quad a_{21}^2 + b_{21}^2 = 1 \quad (18)$$

and are subsequently linearized and introduced to the adjustment solution as weighted constraints

$$-e_s = B_s x + t_s \ ; \ P_s \quad (19)$$

In addition, since the edges to be measured are essentially uni-directional, a linear template edge would perpetually slide along the edge during matching. This effect is compensated by restricting the shift vector of the patch approximately prependicular to the local edge direction as

$$\Delta x\sin\theta + \Delta y\cos\theta = 0 \quad (20)$$

where θ is the angle formed between the edge normal and the y direction [Gruen & Stallmann, 1993]. In matrix form, this condition is expressed as a weighted constraint

$$-e_p = B_p x + t_p \ ; \ P_p \quad (21)$$

The least squares solution for the joint system of equations (14), (19), and (21) is obtained as

$$\hat{x} = (A^T P A + B_p^T P_p B_p + B_s^T P_s B_s)^{-1} \cdot (A^T P l - B_p^T P_p t_p - B_s^T P_s t_s) \quad (22)$$

Iterations are obviously required due to the non-linearity of the mathematical model. From the final updated transformation parameters, an edge point is precisely located in the image as the conjugate of the prespecified edge position in the template.

This edge measurement procedure has been extended into an edge tracking technique, which automatically extracts the complete edge. The user gives an approximate position for the first edge point, the matching algorithm then precisely locates this point and subsequently tracks the edge (Fig. 4). The new approximate match point for the next patch is determined using the previous matched position, its local edge direction and a user-defined incremental distance (in pixels). Edge tracking stops either after the measurement of a prespecified number of edge points or if matching actually fails because the template can no longer find a conjugate window in the image (e.g. the end of an edge or a corner is reached).

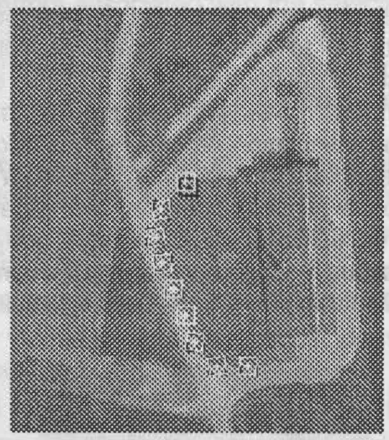


Fig. 4: Edge tracking

Patch sizes typically range between *5x5* and *9x9* pixels. Choice of a smaller patch size results in greater sensitivity to noise, while larger patch sizes increase the probability of interferences by other features close to the current edge.

The significant advantage of the described method lies on the fact that it offers very high positioning accuracies. In addition, its familiar and well established mathematical formulation allows statistical analysis of the results and realistic evaluation of its performance. At the same time though, it is a localized process and as such it is very sensitive to noise and fails in the presence of edge gaps and outline breakpoints.

6. GLOBALLY ENFORCED LEAST SQUARES TEMPLATE MATCHING

The limitations of least squares template matching, due to the use of highly localized radiometric information, can be overcome by its extention into a *globally enforced least squares template matching* strategy (Fig. 5).

An operator initially selects the class to which the object to be extracted belongs (e.g. house, road, land parcel etc.) from an available object class menu and provides manually on-screen approximations of object outline breakpoints, which in essence define a

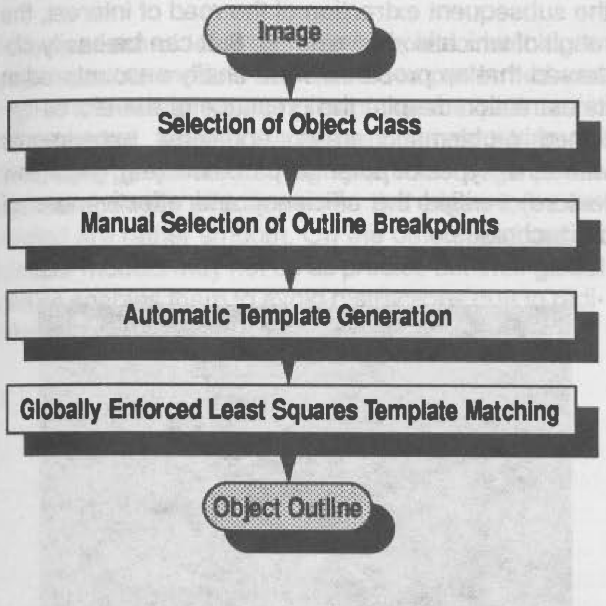


Fig. 5: Globally enforced least squares template matching strategy

polygonic approximation of the edge contour. By interpolating between the polygonic approximation nodes at user-specified intervals, numerous points are extracted, roughly outlining the current object of interest. These points are then used as approximations for the subsequent precise edge positioning through least squares template matching. The synthetic matching template is automatically generated as the 2-D ramp edge which best fits image profiles extracted perpendicular to the polygonic approximation at the interpolated positions. The matching solution, however, will no longer proceed independently for the various points of the same outline. Instead, the mathematical model presented in section 5 is extended to provide a simultaneous global matching solution, whereby globality refers to all matching candidates pertaining to a single object.

More specifically, individual edge positions, detected by least squares template matching along a single object outline, will have to fulfill a certain geometric condition, describing the acceptable outline geometry. For example, it is typical for houses to have edges formed by straight linear segments, while for roads smooth splines are suitable. By defining such a geometric relationship, we are able to tie together the matching adjustment solution of independent points. Assuming straight linear edges for instance, and using the notation of the previous section, it is obvious that edge points along the same line should satisfy the condition

$$\Delta y_i - \Delta x_i \tan\varphi_i = 0 \qquad (23)$$

which in matrix form is expressed as

$$-e_l = B_l X + t_l \; ; \; P_l \qquad (24)$$

Vector X includes the transformation parameters of all matching templates referring to different locations along the same linear edge segment. The associated weights can vary from zero (i.e. constraint virtually removed) to infinity (i.e. constraint strictly enforced). By adding these constraints, the final solution is obtained in analogy to equation (22) as

$$\hat{X} = (A^T P A + B_p^T P_p B_p + B_s^T P_s B_s + B_l^T P_l B_l)^{-1} (A^T P l - B_p^T P_p t_p - B_s^T P_s t_s - B_l^T P_l t_l) \qquad (25)$$

The above matrices include information from all involved templates, and obviously matrices in equation (22) are actually submatrices of their counterparts in equation (25).

For curvilinear features now (e.g. roads), splines are used to ensure piecewise continuity and smoothness. A discrete approximation of this constraint can be given as

$$2\Delta x_i - \Delta x_{i-1} - \Delta x_{i+1} = 0 \qquad (26)$$

$$2\Delta y_i - \Delta y_{i-1} - \Delta y_{i+1} = 0 \qquad (27)$$

$$2\Delta \varphi_i - \Delta \varphi_{i-1} - \Delta \varphi_{i+1} = 0 \qquad (28)$$

Such conditions can be properly fomulated and included as geometric constraints in the global adjustment in lieu of equation (24), and the solution is again given by equation (25). The employment of the appropriate set of constraints is determined by the operator's object class selection at the beginning of the procedure.

By modifying least squares template matching in such manner, its inherent high edge positioning precision and well established mathematical foundation are retained and further improved by the infusion of global geometry information similarly to the snake approach. Thus, the advantages of both techniques are optimally combined providing precise and reliable results.

7. EXPERIMENTS

The semi-automatic strategy for road extraction using wavelet transformed SPOT imagery has been successfully implemented in our Institute on a digital photogrammetric station and Fig. 6, 7 and 8 show a typical example of obtained results.

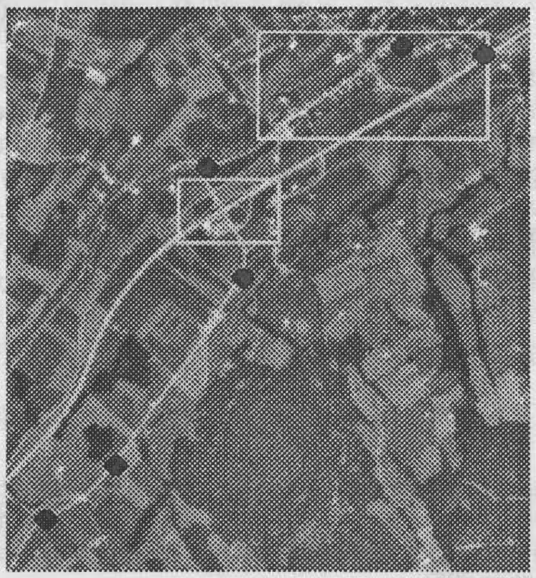


Fig. 6: A portion of a SPOT panchromatic image, with marked seed points and potential problematic areas

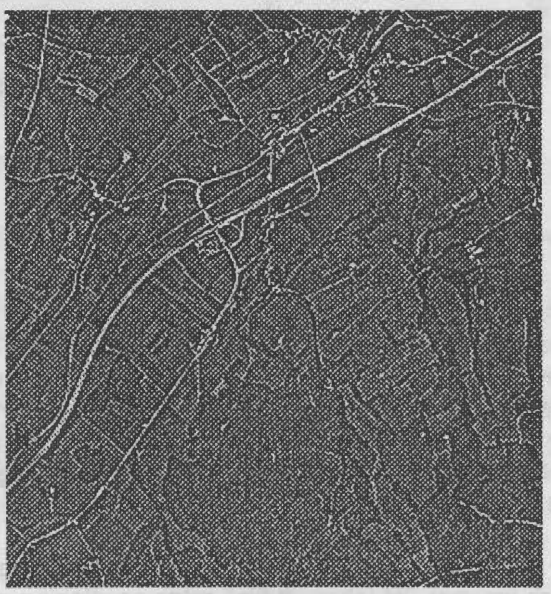


Fig. 7: Wavelet transformed image

Specifically, Fig. 6 depicts a portion of a SPOT panchromatic image of Moudon, Switzerland. On it, marked by black discs, one can distinguish the seed points provided by the operator along the road which is to be extracted. The white rectangles denote potential problematic areas, i.e. intersections with other roads which might cause an erroneous change of direction, and increased radiometric noise and ambiguities due to passage through an urban area. Fig. 7 shows the wavelet transformed image. It is apparent that the transformation has made the existing road network appear more prominently, thus facilitating the subsequent extraction of the road of interest, the result of which is shown in Fig. 8. It can be easily observed that no problems were finally encountered in its extraction despite the presence of the aforementioned problematic areas. Additional experiments with other types of potential problems (e.g. gaps, low texture) verified the efficiency and effectiveness of the technique.

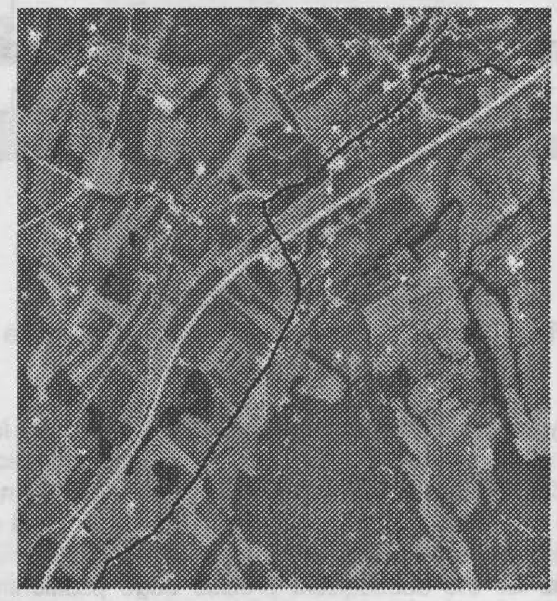


Fig. 8: The semi-automatically extracted road

We have also successfully implemented least squares template matching and Fig. 9 presents at left a typical example of edge tracking by this method. Matching has been initiated at the top of the white dotted line and proceeded downwards. White dots denote the resulting edge positions. Edge tracking stopped at the bottom of the dotted line (position marked by white circle) because it reached a corner, and as previously discussed this caused matching to fail and tracking to stop.

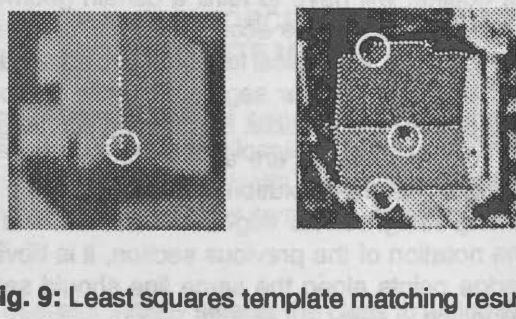


Fig. 9: Least squares template matching results

The right side of Fig. 9 shows the application of this method for the extraction of several edge segments of an object and demonstrates cases of distraction or failure due to its localized character.

A comparison of least squares template matching to active contour models has also been performed and Fig. 10 shows some examples of these experiments. As expected, least squares template matching offered a large amount of highly precise edge points and can be very effective when the edges to be extracted are rather smooth. On the other hand, active contour models may not be as precise but their global nature enables them to avoid distractions due to radiometric or even semantic noise.

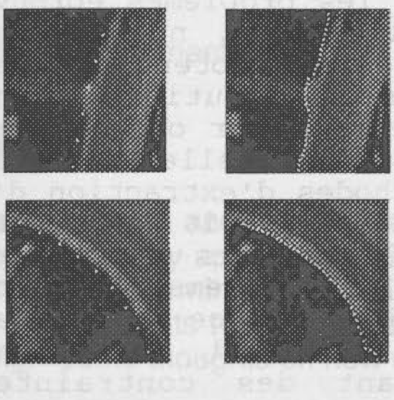


Fig. 10: Edge detection with active contour models (left) and least squares template matching (right)

Globally enforced least squares template matching, presented in section 6 and currently under development, is expected by design to fuse the advantages of both techniques and overcome their inefficiencies.

8. CONCLUDING REMARKS

The semi-automatic object extraction strategies presented in this paper can be ideally used for monoplotting on a digital photogrammetric station. They all have their share of merits and shortcomings which make each of them appropriate for a more or less extensive variety of object types and/or applications. By applying them to orthophotos tied to corresponding DTMs, survey coordinate information can also be assigned to the extracted objects. Thus, a powerful and highly automated digital photogrammetric GIS data capture scheme is emerging, which can substantially improve the performance of this currently cumbersome and relatively costly procedure.

REFERENCES

Ballard D.H. & C.M. Brown (1982): *Computer Vision*, Prentice Hall, Englewood Cliffs, NJ.

Baltsavias E.P. (1991): *Multiphoto Geometrically Constrained Matching*, Mitteilungen Nr. 49, Institute of Geodesy & Photogrammetry, ETH Zurich, Switzerland.

Bartels R.H., J.C. Beatty & B.A. Barsky (1987): *An Introduction to Splines for Use in Computer Graphics and Geometric Modelling*, Morgan Kaufmann Publishers Inc., Los Altos, CA.

Fischler M.A., J.M. Tenenbaum & H.C. Wolf (1981): *Detection of Roads and Linear Structures in Low-Resolution Aerial Imagery Using a Multisource Knowledge Integration Technique*, Computer Graphics & Image Processing, Vol. 15, pp. 201-223.

Fua P. & Y.G. Leclerc (1990): *Model Driven Edge Detection*, Machine Vision & Applications, Vol. 3, pp. 45-56.

Grossmann A. & J. Morlet (1984): *Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape*, SIAM Journal of Mathematics, Vol. 15, pp. 723-736.

Gruen A. & E.P. Baltsavias (1988): *Geometrically Constrained Multiphoto Matching*, Photogrammetric Engineering & Remote Sensing, Vol. 54, No. 5, pp. 633-641.

Gruen A. & D. Stallmann (1992): *High Accuracy Dimensional Measurement Using Non-Targeted Object Features*, International Archives of Photogrammetry & Remote Sensing, Washington, DC, Vol. XXIX, Part B5, pp. 694-700.

Gruen A. & D. Stallmann (1993): *High Accuracy Edge Matching with an Extension of the MPGC-Matching Algorithm*, International Association of Geodesy Symposia, Vol. 108, pp. 339-350.

Kass M., A. Witkin & D. Terzopoulos (1988): *Snakes: Active Contour Models*, International Journal of Computer Vision, Vol. 1, No. 4, pp. 321-331.

Mallat S.G. (1989): *A Theory of Multiresolution Signal Decomposition: The Wavelet Representation*, IEEE Pattern Analysis Machine & Intelligence, Vol. 11, No. 7, pp. 674-693.

Suetens P., P. Fua & A.J. Hanson (1992): *Computational Strategies for Object Recognition*, ACM Computing Surveys, Vol. 24, No. 1, pp. 5-61.

Restitution semi-automatique en mode monoscopique sur un poste de photogrammétrie numérique

Résumé

Cet article porte sur l'extraction semi-automatique d'objets au moyen de l'imagerie numérique en mode monoscopique. Parmi les sujets abordés, on trouve des questions relatives aux concepts, aux algorithmes et à la mise en application, les expériences effectuées, ainsi qu'une discussion sur les problèmes éprouvés et les solutions adoptées. Plus particulièrement, nous présentons l'algorithme et les résultats expérimentaux obtenus concernant l'extraction semi-automatique de réseaux routiers à partir d'images SPOT au moyen d'images transformées par ondelettes. En outre, dans le cas d'images à plus grande échelle et de divers types d'objets, nous comparons deux méthodes d'extraction d'objet que nous avons appliquées et mises à l'essai, soit les techniques du *seuillage par moindres carrés* et des *modèles hypsométriques actifs (Snakes)*. Dans le premier cas, l'emplacement des courbes hypsométriques est déterminé avec précision grâce aux variations locales de l'intensité de gris. Dans le second, on obtient des résultats significatifs en appliquant des contraintes de continuité globale. Finalement, nous proposons une nouvelle approche algorithmique de détection et d'enregistrement semi-automatiques du contour des objets qui combine les avantages des deux méthodes mentionnées ci-dessus. Le seuillage par moindres carrés constitue la base mathématique de cette approche, mais on applique également la continuité globale en introduisant des contraintes de formes en fonction du type d'objet, ce qui garantit des résultats cohérents au plan géométrique.

PERSPECTIVES ON THE USE OF SPATIAL INFORMATION

John D. Bossler

Director, Center for Mapping, The Ohio State University
1216 Kinnear Rd., Columbus, Ohio 43212, United States of America

ISPRS Commission II

KEY WORDS:

Spatial data, Elements of Usage, Accuracy, Standards, Sharing.

ABSTRACT

Accurate, current and well-defined spatially referenced digital data is in demand to satisfy the hungry appetities of GIS. The data most in demand is listed and some comments concerning the required accuracy of these data are related. Several major impediments to sharing data are discussed along with incentives to enhance sharing. Some thoughts on new ways of acquiring digital are provided.

INTRODUCTION

The title of this paper, "Perspectives on the Use of Spatial Information," allows me to discuss a wide array of topics that are important in the information age. Since both this paper and the presentation of it are finite in length, I have chosen topics that are on my mind as a participant in today's GIS/LIS community.

In an overall sense, the use of GIS is increasing at an enormous rate (Daratech, 1993) because data is becoming available, the systems are becoming easier to use and most importantly, because the use of modern GIS enables users to make important decisions quicker and more economically than by traditional means. On the other hand, the use of digital spatial information is impeded by many factors. Some of the most important impediments are the lack of digital data itself, understanding (diffusion) of the technology, training, lack of information related policy, standards, incentives, quality control/quality assurance methods and institutional problems.

The discussion in this paper is directed toward identifying the data and information most in demand, and identifying a subset of the impediments to acquiring, sharing and using these data.

NEEDED DATA AND USAGE

From a national perspective in the United States, a number of data sets are being created and used at an almost insatiable rate. It is difficult to prioritize these and even more difficult to develop a taxonomy for them, however, it seems clear that one data set is unique and fundamental to all the others--that data set is geodetic control. Geodetic control provides the spatial component to all other data sets and hence, can be considered to be the most important. Digital terrain data and digital imagery (Mapping Science Committee, 1994) are considered to be fundamental for GIS usage, especially at the national level. In the above report, the Mapping Science Committee (MSC) has identified five other important data sets:

1.) The digital orthophoto

2.) A street centerline file

3.) A cadastral framework

4.) A natural resource framework

5.) A hydrological framework

A recent user survey by the Ohio Geographically Referenced Information Program (OGRIP) confirms these selections at the "local" level (Ohio Geographically Referenced Information Program, 1993). In this survey the following data elements were considered to be of the greatest interest:

1.) Geodetic control

2.) Parcel attributes

3.) Right-of-Ways

4.) Bridges

5.) Parcel boundaries

6.) Street attributes

7.) Street centerline

8.) Municipal boundaries

At first glance, the difference between national needs and the local needs appears to be in the natural resource area, however, the OGRIP survey found that the top five data sets that *needed* to be entered into the respondents databases were:

1.) Flood plain

2.) Watersheds

3.) FEMA zones

4.) Land use

5.) Zoning

It is important to note that the combination of "most interest" and "most needed" data by the OGRIP survey, corresponds to the more national findings of the mapping science committee.

ACCURACY CONSIDERATIONS

The OGRIP survey was weak concerning accuracy requirements however according to that survey, almost all "GIS" accuracy requirements were satisfied by $s = \pm 2.5$ feet. It is pointed out by the MSC that one man's "accurate" is another "inaccurate" depending on the application. The following table illustrates that fact. In any event and contrary, I might add, to recent emphasis by the National Geodetic Survey, most applications of spatial data leading to decisions relating to land development, environmental problems and even utility line siting, are satisfied by accuracies of $s = \pm 0.1$m. This assertion differentiates between surveying and GIS usage. However, it should be pointed out that in a recent article (Fernandez-Falcon, E., et. al., 1993.) a suggestion is made that spatial accuracy needs for surveying may actually be declining, except for geodynamics, because error propagation resulting from GPS surveys is completely different from that encountered by traditional techniques, and the cost per point has

Accuracy	Example Applications
0.001 m	Crustal motion, geodynamics, geophysics
0.01m	Property surveying, civil engineering
0.1m	Cadastral mapping
1 m	Facilities management for utilities
10 m	Mapping, soil and wetland mapping
100 m	Small scale mapping, National Biological Survey
1000 m	Ice flows, global change research

Table 1: Example applications of spatial data and their corresponding positional accuracy requirements.

been dramatically reduced. It appears therefore, that the accuracy of spatial data necessary to satisfy growing market demands is satisfied by technology like that developed at The Ohio State University Center for Mapping (Bossler, J., et. al., 1994.)

IMPEDIMENTS TO USAGE

Standards

The need for data standards in order to effectively use and share data has long been recognized (Mapping Science Committee, 1990.) The most needed standards are content, accuracy and transfer standards. However, the discussion below focuses on transfer standards.

Federal Responsibility. While it is arguable, there is general agreement that standards are a federal responsibility. However, there still seems to be a reluctance to accept this idea by some federal, private, and professional organizations. It is important to note at the onset of this argument that the top down (federal) development of standards has existed for many years, e.g., in the mapping arena the 7 1/2' quadrangle is a (de facto?) standards.

I believe that setting standards may well be the most important thing a federal agency could do because setting standards should be done by an objective, unbiased, public organization so that no one private firm is unfairly favored. Another characteristic of an organization that sets standards is that it should be inherently the organization of "ultimate" responsibility, e.g., the United States Geological Survey (USGS) is the ultimate mapping authority in the U.S.

A spatial data standard should reach as many practitioners as possible, be widely available and have the potential--at least--of being maintained, changed, and updated until it is obsolete. Federal agencies have the mandate, scope and visibility to accomplish this development. These points confirm the assertion that spatial data standards should be developed and maintained by federal agencies. It goes without saying that this should be done in cooperation with many other organizations by soliciting their views and supplying funding.

Research. Presently there is virtually no enthusiasm in the GIS community for the idea that universities--or for that matter federal agencies--should perform research that fosters the development of standards. Research on standards was an explicit recommendation of the MSC in a recent report (Mapping Science Committee, 1990) to the USGS, but I am unaware of *any* federal research grants or contracts related to standards. Some of the questions that need research are:

What standards should be developed and in what priority?

How can we evaluate their efficiency?

What are the benefit/cost ratios?

What conceptual data model should be used for what standard?

What primitives should be used for various data models and types?

How are standards updated, maintained and expanded?

What process should be used to develop or test a standard?Some of these questions are deceptively simple. I challenge my colleagues to consider them carefully and try to provide us with answers. I am fully aware of the fact that the USGS developed Spatial Data Transfer Standard (SDTS) was conceived primarily by academicians. However, at the moment I know of very little research related to this important aspect of data sharing. Moreover, the fact that SDTS will have taken 13 years to emerge as a FIPS standard reinforces the need for the above. The need for these activities is becoming clearer and has recently been documented (Morrison 1988).

Sustained Support. The development and promulgation of an important, complex standard requires a significant outlay of resources over a long period of time. This fact seems not to be apparent a priori to many organizations that have developed standards. Important aspects of the undertaking include initial testing, user manuals and guides that include clear examples, compliance testing, a service support organization and a maintenance function. Moreover, in order to accomplish widespread acceptance and utilization of the standard, "marketing' and "advertising" efforts must be organized and supported. Public hearings at appropriate meetings are essential. These activities cost a lot of money and require the best people in any given organization since, for example, the compliance testing function requires a complete and thorough knowledge of the entire standard. The requirement for these types of activities needs to be acknowledged at the beginning of the standards development process.

ACCELERATING USAGE

We need to provide solutions to the problems inhibiting the use of spatial data by developing the appropriate standards policies, etc. The most important problem is that of obtaining the digital data. However, this means acquiring the data and making it accessible to all interested individuals or agencies. In an attempt to contain the length of this paper, I have omitted any discussion of the information highway and assume it will be covered elsewhere. Moreover, the most frequently used method of acquiring digital spatial data is to digitize source (map) data.

New Paradigm

I will argue that modern mobile mapping systems are highly competitive means of acquiring digital data. Clearly, with two wavelength GPS technology (Dedes, G., Goad, C., 1994) coupled with modern digital cameras (Bossler, et. al., 1994) accuracies of $s = \pm 0.1$ meters are possible from mobile platforms traveling at normal speeds. For the purposes of this discussion, I include airplanes, boats, highway and railroad vehicles to be "mobile" vehicles.

Therefore, since the accuracy requirements are satisfied by such mobile platforms there is only the question of economics.

Locating a feature (e.g., an electric pole) to ± 0. 1 m. by traditional surveying techniques costs between five dollars and two hundred dollars depending on factors such as numbers of features, length of traverse, distance to reference control, etc. With mobile mapping systems an "apple to apple" comparison yields reductions in cost of factors of two to fifty. Therefore, I assert that at least *consideration* should be given to directly acquiring *current* accurate digital data rather than digitizing old inaccurate maps. The use of digital cameras provides enormous additional benefits such as the ability to keep the data in a database, retrieve it at will and measure something that was not interesting at the time the project was accomplished.

The applications of this history file in digital form, are limited primarily by our imagination.

Incentives for Sharing Data

A number of incentives for data sharing and other forms of cooperation, which will obviously enhance usage, appear to have worked--and in some cases very well. For example, Richard Yorczych of the Horizontal Network Branch of the National Geodetic Survey (NGS) reports that they have received data from other organizations for 65,000 horizontal geodetic control points since 1980 (Yorczych, 1992). This data donor program works because the donors (state, county, and private organizations) want to assure the accuracy of the points they observed and earn a stamp of approval from the nation's highest authority on geodetic control. It was essential that the NGS develop a standard in order to implement this mechanism. This standard for submitting geodetic information to NGS is known in the vernacular as the "Blue Book" (U.S. Department of Commerce 1980, U.S. Department of Commerce 1988). The cost of these data can be conservatively estimated at 65,000 x $1,000/point = $65,000,000. The *value* of such data was the subject of a study by Epstein and Duchesneau, 1984, and while cost and value are not always equal, they are generally positively correlated.

One of the best examples of a successful incentive program at the state level can be found in North Carolina (Holloway, 1986). Ten years ago, North Carolina began providing seed money to counties that followed certain mapping and other standards developed by the state. These standards in turn comply with federal standards. As far as I know, every county in North Carolina has taken advantage of this program. Since the amount of funding given to each county is small, it appears that money may not be the only motivating factor, and that other less tangible factors like those mentioned in the preceding paragraph are important.

CONCLUSIONS

Spatial data and information are in demand and are being increasingly used to satisfy the hungry appetites of GIS. The most important data sets are, apparently geodetic control, digital orthophoto, street centerlines, parcel boundaries, land use and hydrological data. Accuracy and quality of data is important but should be viewed from a user of spatial data perspective not from say, a photogrammetrist's perspective. Standards, QA/QC procedures, new data acquisition methods, and incentives for sharing, should be topics of research, development and the implementation of them should go on in parallel, as soon as possible.

Non-monetary incentives may be able to be used to foster institutional collaboration at the local, state and national levels. Consideration should be given to acquiring digital spatial data and information directly from field observations rather than digitizing out of date maps.

It is an exciting time in the field of surveying and mapping and change is the order of the day. While some technologies will disappear, other more powerful tools will be adopted and overall a higher quality product will emerge to better serve society.

REFERENCES

Bossler, J., Novak, K., and Johnson, P., 1994. Digital Mapping on the Ground and from the Air. In: GeoInfo Systems, January, pp. 44-48.

Daratech, Inc., 1993. Geographic Information Systems Markets and Opportunities. Prospectus, Cambridge, Massachusetts, USA.

Dedes, G., Goad, C., 1994. Real-time cm-level GPS Positioning of Cutting Blade and Earth Moving Equipment. In: National Technical Meeting Proceedings, The Institute of Navigation, Alexandria, VA-USA.

Fernandez-Falcon, E., Strittholt, J.R., Alobaida, A.I., Schmidley, R.W., Bossler, J.D. and Ramirez, J.R., 1993. A Review of Digital Geographic Information Standards for the State/Local User. In: URISA: Journal of the Urban and Regional Information Systems Association, Vol. 5, Number 2, pp. 21-28.

Holloway, D., 1986. Land Records Management in North Carolina. In: Popular Government, Institute of Government, University of North Carolina, Chapel Hill, NC-USA.

Mapping Science Committee, 1990. Spatial Data Needs: The Future of the National Mapping Program. National Academy Press, Washington, D.C.-USA.

Mapping Science Committee, 1994. National Spatial Data Infrastructure: Selected Findings and Data Sets. In press. National Academy of Sciences/National Research Council, Washington, D.C.-USA.

Morrison, J., 1988. Digital Cartographic Data Standards: The U.S. Experience. In: Cartography, Issue 1712.

Ohio Geographically Referenced Information Program, 1993. Survey of Users of Geographic Information Management Systems, Comprehensive Report, Columbus, OH-USA.

Yorczyk, R., 1992. Telephone conversation with author, 4, February.

CONSIDÉRATIONS SUR L'EMPLOI DES DONNÉES À REFÉRENCE SPATIALE

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif

JOURNEY FROM GIS IGNORANCE TO IMPLEMENTATION - A TRUE STORYU FROM THE MIDDLE EAST

Zul Jiwani
The Centre for GIS, State of Qatar

Abstract

Not available at time of printing

La construction d'un SIG à partir de zéro - Cas vécu au Moyen-Orient

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif

THE USE OF GPS FOR GIS GEOREFERENCING: STATUS AND APPLICATIONS

M. Elizabeth Cannon
Department of Geomatics Engineering, The University of Calgary,
2500 University Drive N.W., Calgary, Alberta, T3H 1Z2, Canada

KEY WORDS: GPS, GIS, surveying, georeferencing, receiver, static, kinematic

ABSTRACT

The use of GPS for georeferencing has entered a new age with the operational capability of GPS system as well the developments in GPS receiver technology and processing strategies. The current status of GPS methods and accuracies are reviewed in the context of GIS applications. Trends in GPS hardware are also presented. Several applications of GPS for the acquisition of spatially-related data are discussed. These include applications in the agricultural sector for salinity mapping and variable fertilizer spreading. Other applications in urban environments are also presented where shading problems may limit the achievable accuracy and availability of GPS. The paper concludes with a discussion of the trend in GPS processing developments, especially high precision real-time positioning, and its impact in the GIS field.

1. INTRODUCTION

The Global Positioning System (GPS) has already made a major impact into the fields of surveying and georeferencing. With the completion of the satellite constellation, the reduction in receiver costs, as well as improvements in processing algorithms and related software, the growth in applications is expected to continue. Most recently, the GIS field is embracing the use of GPS for moving platform applications, the so-called kinematic approach where large amounts of data can be collected in a cost-effective manner. The following papers focuses on the current state of the GPS system, the options for receiver hardware selection, the numerous modes of data collection as well as a cross-section of kinematic GPS applications for the land, marine and airborne environments.

2. STATUS OF GPS AND TRENDS IN RECEIVER TECHNOLOGY

GPS is entering a new era now that the system is nearing completion and moving to the operational phase. There are currently 26 satellites in orbit of which 3 are prototype Block I. Under this constellation, the geometry is effectively consistent throughout most of the day as evidenced by the Position Dilution of Precision (PDOP) which is given in Figure 1 for satellites in view at Calgary in March, 1994. The PDOP is a figure of merit for the satellite geometry and a value below three can be considered good. Since the satellite coverage is generally below three for most of the day, detailed mission planning is not as important as in the past except when satellites become unhealthy which may cause outages. Another Block II satellite is expected to become operational in mid-March, 1994 which will further enhance the current coverage.

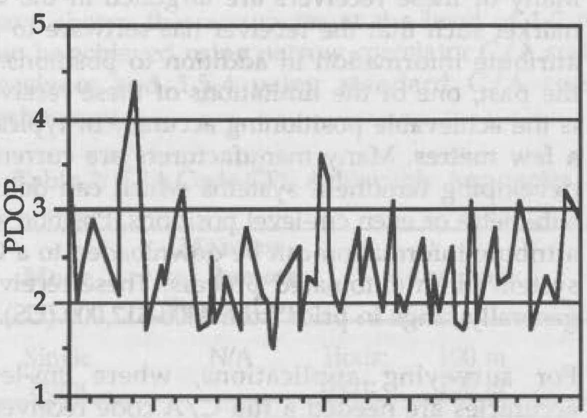


Fig. 1: Position Dilution of Precision (PDOP) for Calgary in March, 1994

With the availability of continuous GPS coverage, the use of the system for a wide range of applications is increasing at high rate. In order to meet these range of applications, GPS receiver manufacturers currently offer a wide range of products that vary in size, cost, performance and capabilities. Table 1 gives a list of some of major classes of GPS receivers. The first is the Coarse/Acquisition (C/A) code engine which is a GPS receiver on a circuit board having dimensions as small as 5 cm x 5 cm. These receivers are relatively new on the market and are ideal for system integration applications where GPS is merely a component in the overall system. Most of these receivers have 5 to 8 channels which means that all the satellites which are in view may not be tracked. Receivers use algorithms based on the PDOP or the satellite elevation in order to select satellites. Tests have shown that the algorithm may affect the receiver's performance under shading conditions (McLellan et al., 1994). Another distinguishing factor regarding GPS engines is the availability of raw data to the user. Some receivers do not output the measurements, but only the internally computed positions, whereas other receivers output both the pseudorange and carrier phase data. This feature is important for system developers since raw data is usually required. The cost of the C/A code engine class of receivers is generally between $500 - $8,000 (US), depending on the performance. These receivers will play a major role in the GIS market since they can add a georeferencing component to a GIS system at relatively low cost.

The second class of the receivers is the C/A code handheld type which are effectively GPS engines housed in a small, lightweight data collector. Many of these receivers are targetted at the GIS market such that the receiver has software to log attribute information in addition to positions. In the past, one of the limitations of these receivers is the achievable positioning accuracy of typically a few metres. Many manufacturers are currently developing handheld systems which can deliver sub-metre or even cm-level positions. Position and attribute information can be downloaded to a GIS system in an automated process. These receivers generally range in price from $800-$12,000 (US).

For surveying applications, where cm-level accuracies are needed a full C/A code receiver is generally used. These receivers have a sophisticated user interface which allows more flexibility in the operation of the unit and the number of channels range from 8 to 12 so that all satellites in view may be tracked. Raw data is generally available to the user which may be processed by manufacturer-supplied software or alternatively by third party programs. The quality of the receiver oscillator as well as the antenna are generally higher than either the engine or handheld receivers, so that errors (i.e. measurement noise and antenna phase centre stability) will be minimized. These receivers range in cost from $12,000 to $20,000 (US) and may be used for high-end GIS applications where the achievable accuracy is important. Alternatively, these receivers are often used for differential base stations in GIS applications since they have the all-in-view capability.

The final class of receiver is the so-called P codeless unit which gives dual frequency carrier phase (and possibly pseudorange) data without access to the encrypted P code (i.e. the Y code). These receivers are the current state-of-the-art and are generally priced in the 35,000-45,000 (US) range. The main applications of these receivers are for high accuracy rapid static surveying, on-the-fly kinematic surveying (OTF) as well as high precision static surveying over long baseline lengths (say >50 km).

Table 1: GPS Receiver Classes and Features

Class	General Features
C/A code engine	- low cost - 5-10 channels - system integration required - raw data may not be available - flexible architecture - manufacturer post-processing software generally not available
Handheld C/A code receiver	- low-mid cost - 5 or 6 channels - may have manufacturer supplied software for m-level accuracies - may interface to GIS package
C/A code receiver	- mid-range cost - 8-12 channels - sophisticated user interface - manufacturer supplied software for cm-level accuracies
P codeless receiver	- premium cost - 9-12 channels - sophisticated user interface - complete system with manufacturer-supplied post-processing software - dual frequency data available after GPS becomes operational - quality of L2 data may vary

Receiver hardware costs are reducing at a fast rate which is mainly driven by the increasing use of the GPS system. An interesting phenomenon that is occurring is that the spread in the cost, i.e. from the lowest to the most expensive, is growing wider. This is due to the emergence of GPS engines which are 'bare-bones' receivers as compared to the full turn-key systems that include software and extensive customer support. Figure 2 shows the trend in GPS hardware over the past several years, after McDonald (1989).

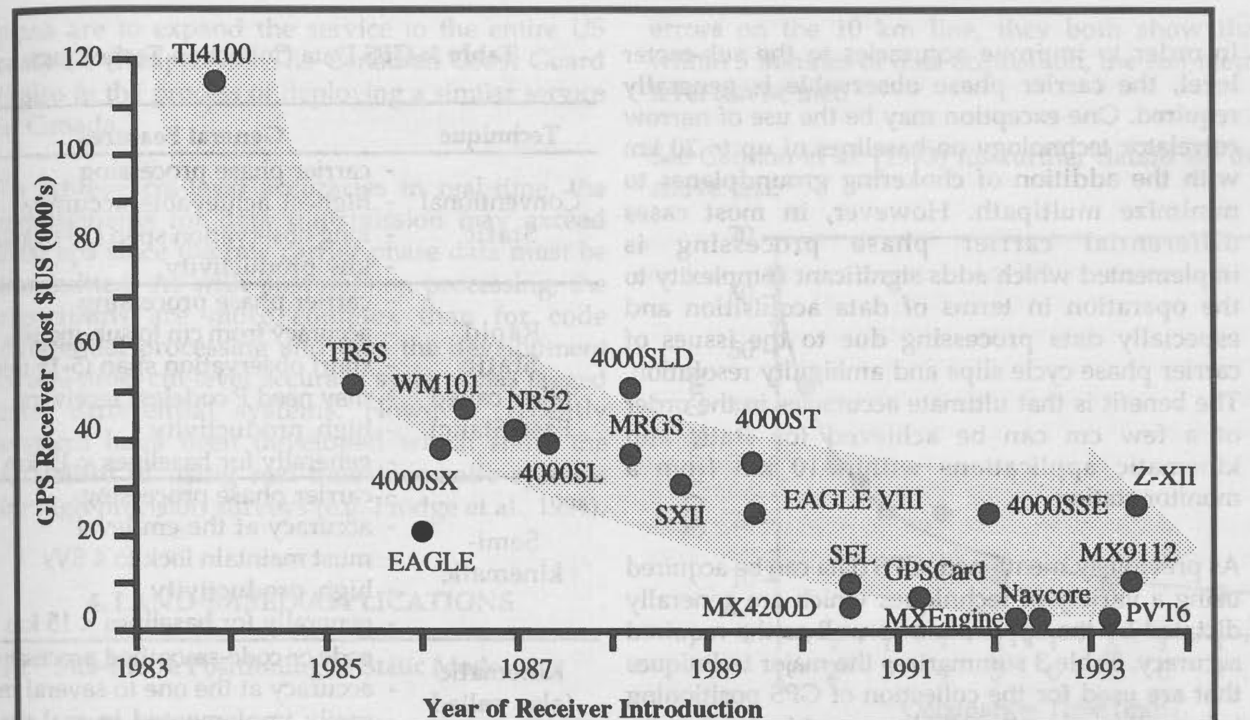


Fig. 2: GPS Receiver Cost Trend

3. ACHIEVABLE ACCURACIES AND OPERATING MODES

Achievable accuracies using GPS are generally dependent on the type of receiver selected, the measurement utilized (i.e. code versus carrier phase, single point versus differential), as well as the dynamic environment (i.e. static versus kinematic land, marine or airborne). This paper is not intended to overview all the above in detail, however the major accuracies are summarized in Table 2.

The table is sub-divided between single point and differential modes. Single point positioning is limited by the GPS system itself and is at the level of 100 m horizontally and 156 m vertically. Due to this limitation, single point positioning is not dependent on the receiver used, assuming C/A code technology. Differential accuracies are determined by the type of observable used, i.e. code (pseudorange), carrier phase, or a combination of both. For the case when the code measurements are used, accuracies are at the level of 2-9 m horizontally and 3-11 m vertically. The lower end of the scale is for the new generation of narrow correlator C/A code receivers which have a C/A code resolution of 10-20 cm (e.g. Van Dierendonck and Fenton, 1992) whereas the higher end is for standard C/A code receivers. These accuracies can be achieved in static or kinematic mode.

With the addition of carrier smoothing of the code, the achievable accuracy is improved. Tests have shown that accuracies at the level of 1-2 m can be achieved using narrow correlator C/A code receivers and 1.5-4 using standard C/A code technology.

Table 2: C/A Code GPS Achievable Accuracies

Mode	Monitor to Remote Separation	Accuracy	
Single Point	N/A	Horiz: Vert:	100 m 156 m
Differential Code	50 km	Horiz: Vert:	2 to 9 m 3 to 11 m
	500 km	Horiz: Vert:	5 to 11 m 6 to 12 m
Differential Smoothed Code	50 km	Horiz: Vert:	1 to 3 m 1 to 4 m
	500 km	Horiz: Vert:	3 to 9 m 4 to 10 m
Differential Carrier Phase	50 km	Horiz: Vert:	0.05 to 0.3 m 0.1 to 0.5 m
	500 km	Horiz: Vert:	1 to 4 m 2 to 6 m

In order to improve accuracies to the sub-meter level, the carrier phase observable is generally required. One exception may be the use of narrow correlator technology on baselines of up to 20 km with the addition of chokering groundplanes to minimize multipath. However, in most cases differential carrier phase processing is implemented which adds significant complexity to the operation in terms of data acquisition and especially data processing due to the issues of carrier phase cycle slips and ambiguity resolution. The benefit is that ultimate accuracies in the order of a few cm can be achieved for static and kinematic applications within 10 km from a monitor station.

As previously mentioned, GPS data can be acquired using a variety of techniques which are generally dictated by the application as well as the required accuracy. Table 3 summarizes the major techniques that are used for the collection of GPS positioning information as well as their general features.

For static applications there are generally three techniques which are used, namely conventional static, rapid static and semi-kinematic. The first technique is the most robust and reliably gives results at the 1-3 ppm level (i.e. 1-3 cm per 10 km of monitor-remote separation). Its main disadvantage is the low productivity caused by the long observation span. In recent years rapid static surveying has become feasible due to the increased satellite constellation, P codeless receivers as well as improved processing algorithms. In general this technique can give cm-level accuracies in minutes, but is limited to baselines where integer ambiguities can be resolved, i.e. generally less than about 15 km. If rapid static surveying is implemented using a receiver that does not output the carrier phase (only the carrier-smoothed code), the accuracy is limited to about 1-3 m. Semi-kinematic positioning is an alternative to rapid static in that it does not depend on P codeless technology. In this case the carrier phase integer ambiguities are initialized at a starting point and then the receiver is transported to other points of interest. As long as lock is maintained to at least four satellites during these movements, cm-level accuracies can be achieved. However, if less than four satellites are tracked, the ambiguities must be re-initialized. It is expected that the rapid static technique will fully replace semi-kinematic positioning once the cost of P codeless technology decreases further. For more information on semi-kinematic surveying see Cannon (1991).

Table 3: GPS Data Collection Techniques

Technique	General Features
Conventional Static	- carrier phase processing - highest achievable accuracy - long observation span (30-60 min) - low productivity
Rapid Static (also called Fast Static)	- carrier phase processing - accuracy from cm to sub-metre - short observation span (5-15 min) - may need P codeless receivers - high productivity - generally for baselines < 15 km
Semi-kinematic	- carrier phase processing - accuracy at the cm-level - must maintain lock to 4 SVs - high productivity - generally for baselines < 15 km
Kinematic (also called Dynamic)	- code or code-smoothed processing - accuracy at the one to several m - easily implemented in real-time - high productivity
Kinematic with Static Initialization	- carrier phase processing - static initialization required - accuracy at the several cm level - generally for baselines < 50 km - must maintain lock to 4 SVs - high productivity
On-the-fly Kinematic (OTF)	- carrier phase processing - no static initialization required - accuracy at the several cm level - generally for baselines < 50 km - generally need P codeless units - high productivity

The need for real-time positioning is usually driven by navigation or guidance requirements, however knowledge of high quality results during the observation span is also beneficial. Of the techniques presented in Table 3, the kinematic (dynamic) mode of operation is the easiest to implement in real-time. This is due to the use of the code observable which simplifies the processing requirements as wells as data transmission. Corrections from the monitor receiver can be transmitted at a rate as low as 50 bps using an internationally accepted standard format, i.e. RTCM-SC104. Accuracies in the real-time mode of operation are generally within those achieved in post-mission, assuming no significant breaks in the correction transmission. The U.S Coast Guard has initiated differential GPS (DGPS) correction transmission from its existing marine radiobeacons and the current coverage includes part of the US east coast as well as the Gulf of Mexico. Future plans are to expand the service to the entire US coastline (Alsip,1993). The Canadian Coast Guard is also in the process of deploying a similar service in Canada.

To achieve cm-level accuracies in real-time, the requirements for data transmission may exceed 2000 bps since the raw carrier phase data must be transmitted. As with post-mission processing, the algorithms are more complex than for code differential processing and thus the development of real-time, cm-level accuracy systems has lagged code differential systems. However, recently systems have been developed which show the feasibility of using real-time kinematic systems for high precision surveys (e.g. Frodge et al., 1994).

4. LAND-BASED APPLICATIONS

4.1 Sub-Metre Positioning in Static Mode

Many GIS applications traditionally required a few metre accuracy for georeferencing static points occupied during a survey. The current emphasis is on sub-metre accuracy which can place particular constraints on the type of GPS receiver technology that must be used as well on the time required to occupy a point. Generally, the GIS community uses standard C/A code technology which delivers 1-3 m accuracy. In order to improve the accuracy below a metre with these receivers, the carrier phase observable must be used and the point occupied for a longer period of time in order to acquire sufficient satellite geometry.

A study to determine the time required to reach the sub-metre level, a test was conducted with the Motorola LGT1000™, a GPS/GIS terminal which can track six satellites simultaneously and can also output the carrier phase measurement. Data was collected on baselines of 500 m and 10 km. About 2 hours of data were recorded for each baseline, and post-processed using The University of Calgary's SEMIKIN™ program (Cannon,1990). The data was processed in subsets and the resulting coordinates were compared to the known baseline coordinates in order to determine the achievable accuracy. Figure 3 shows the relationship between site occupation time and 3-D accuracy for the 100 m baseline. Figure 4 shows the same for the 10 km baseline.

The two figures clearly show that the achievable accuracy improves as a function of the site occupation time. Although the results are slightly different for the two baselines due to the increased errors on the 10 km line, they both show that within 5 minutes of data acquisition, the sub-metre level can be met.

See Cannon et al. (1993) for further details on the above test.

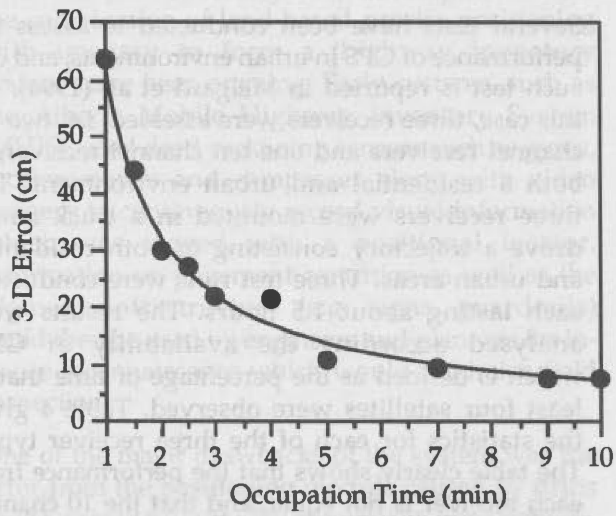


Fig. 3: Position Accuracy as a Function of Station Occupation Time - 500 m Baseline

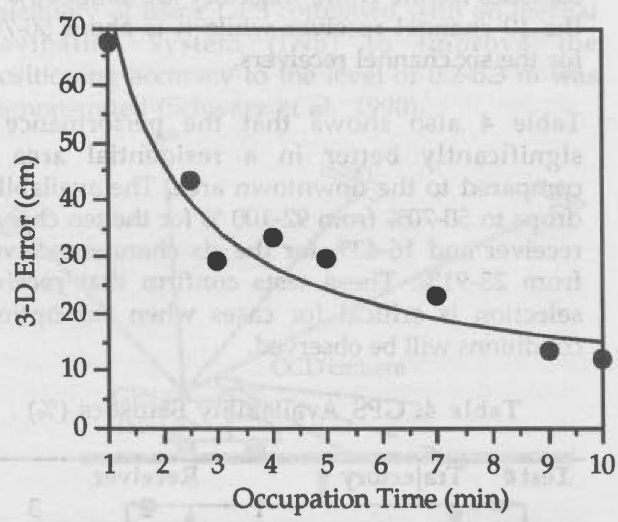


Fig. 4: Position Accuracy as a Function of Station Occupation Time - 10 km Baseline

4.2 Use of GPS in an Urban Environment

One of the major limitations of using GPS for GIS applications is the shading problem that is experienced under foliage and near buildings. Although it is an important concern for static applications, it is magnified for kinematic applications when a continuous trajectory is usually required. The susceptibility of GPS to shading is partially a function of the receiver that is used. For example, it is expected that a 12 channel receiver will generally be less susceptible to multipath compared to a 6 channel receiver. Also, differences in the receiver tracking loop design may also affect the performance of a particular receiver in a sub-optimal environment since this may affect both multipath susceptibility and satellite re-acquisition times.

Several tests have been conducted to assess the performance of GPS in urban environments, and one such test is reported in Melgard et al. (1994). In this case, three receivers were assessed, i.e. two six channel receivers and one ten channel receiver, in both a residential and urban environment. The three receivers were mounted in a truck which drove a trajectory consisting of both residential and urban areas. Three test runs were conducted, each lasting about 1.5 hours. The results were analysed according the availability of GPS, which is defined as the percentage of time that at least four satellites were observed. Table 4 gives the statistics for each of the three receiver types. The table clearly shows that the performance from each receiver is not equal, and that the 10 channel receiver has better availability statistics than the six channel receivers, as expected. The statistics for the entire trajectory are about 90% for the 10 channel receiver while it is about 50-70% for the six channel receivers.

Table 4 also shows that the performance is significantly better in a residential area as compared to the downtown area. The availability drops to 50-70% from 92-100 % for the ten channel receiver and 16-43% for the six channel receivers from 23-91%. These tests confirm that receiver selection is critical for cases when sub-optimal conditions will be observed.

Table 4: GPS Availability Statistics (%)

Test #	Trajectory	Receiver 1	Receiver 2	Receiver 3
1	Downtown	51	16	40
	Residential	98	91	78
	Entire Traj.	88	67	73
2	Downtown	69	39	43
	Residential	100	68	61
	Entire Traj.	92	66	75
3	Downtown	57	31	35
	Residential	92	24	23
	Entire Traj.	88	51	60

4.3 Precision Farming

The application of GPS to the agricultural industry is beginning to accelerate due to the reduction in receiver costs. The University of Calgary, together with Alberta Agriculture, is currently developing a system which can be used to measure the variability of crop production as a function of location (Gehue et al., 1994). The system consists of several components which include GPS receivers for positioning, a yield monitoring system which outputs the instantaneous Bu per acre, an EM conductivity meter for salinity measurements, and soil samples for determination of soil types and nutrients. The DGPS/yield monitoring data can be collected under normal combining operations. Two 10-channel C/A code narrow correlator spacing NovAtel GPSCard™ sensors were used as they have shown to provide sub-metre accuracy in previous field tests using a robust carrier phase smoothing of the code approach (e.g. Cannon and Lachapelle, 1992). Once all this data has been collected and analysed, variable fertilizing can be used in subsequent years to optimize productivity. DGPS can be used as a real-time navigation and guidance tool for this application.

A GIS is used to organize and collate the variety of data collected from year to year. Various layers of information can be draped (using surfacing routines) over each other through the DGPS position. It is then possible to overlay crop yields with salinity, for example, and draw conclusions based on the relationship (i.e., high salinity results in low yield). Further, predictions can be made on how to handle these regions when it is time to fertilize and/or seed. The GIS can also be updated with external information such as aerial photography or remote sensing. Current information layers of topography, salinity, soil type, nutrients, and crop yields were collected for the 1993 harvest.

The goal of the precision farming project is to optimize the yield-input relationship for any given field. Since this is the first harvest season of the project, results to date only give an indication of the variation within the field for this year, with the condition that it has been treated equally in the past. A surfacing routine is used to view the yield variations with the result shown in Figure 5 for a subset of the test site. The figure clearly shows that the field is not homogeneous in yield. Surface analysis of crop yields allows agricultural researchers to identify problem areas and to maintain data quality by detecting results such as the sudden peak in the NE corner of the field which is most likely caused by high moisture in the swath which returns a false reading from the yield sensor.

Fig. 5: Yield Map (3rd Dimension is Yield)
Hussar Site Subset

Once all of the necessary data has been sorted into their respective layers, relationships and effects between each layer can be determined. This information can be used to optimize the field potential by treating the field based on the specific sub-class of different sections. A field may have several distinct classes of soil that should have different quantities and mixtures of fertilizers applied to it in order to get maximum yield. The variable rate application would also take into account other variables such as salinity, topography and history of previous crops and applications.

The test field consists of gently rolling hills with a steep north facing hill in the middle. The GPS monitor station was installed near the field and the moving platform was operating within a few km from the reference station. The crop was harvested on September 20 and 21 1993. On November 9, soil samples were taken at various locations for cross-referencing the first dataset.

Table 5: RMS Agreement Between Carrier Phase Smoothing and OTF Solutions at Crossover Points

Date	RMS of Differences		
	East (m)	North (m)	Height (m)
Sept. 21	0.14	0.21	0.51
Nov. 9	0.14	0.26	0.66

The positioning accuracy requirements in this project are 0.5 m horizontally and 1.0 m vertically. Two techniques were used to reduce the DGPS data, namely a carrier-smoothing of the code technique, and a on-the-fly ambiguity resolution procedure (OTF). The achievable height accuracy was verified by comparing the estimated positions using these two techniques, i.e. the OTF solution was used as a reference trajectory. Table 5 shows the results for each of the two test days and illustrates that the positioning requirements are being met using the current configuration.

4.4 Highway Inventory

Over the past several years, developments into the combination of land-based precise positioning with imagery to form a 'highway inventory system' have been ongoing. Early systems, such as the Alberta Mobile Highway Inventory System (MHIS), used dead-reckoning sensors such as gyros, accelerometers and compasses along with video imagery to continuously record visual information which was tagged with a positional locator. Information on pavement condition as well as the highway infrastructure (e.g. signs, guardrails) could then be used by engineers and planners for in-house reconnaissance which would minimize field inspections.

One of the major drawbacks of the system was the error from the positioning module caused by drifts in the sensor output. To circumvent this problem, the sensors required frequent calibration which greatly affected productivity. In 1988, the feasibility of using GPS combined with an inertial navigation system (INS) to improve the positioning accuracy to the level of 0.2-0.3 m was demonstrated (Schwarz et al. ,1990).

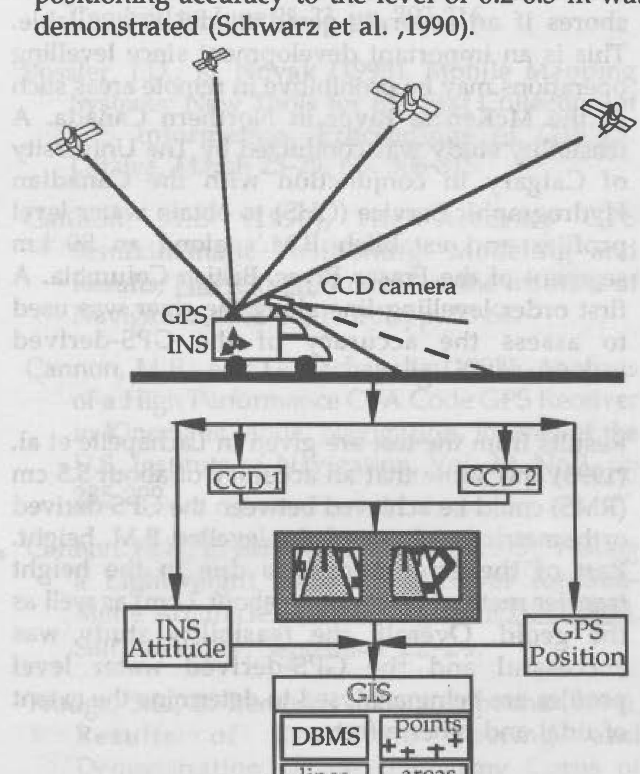


Fig.6: VISAT Concept (Schwarz et al.,1993)

A new initiative at The University of Calgary is the VISAT program which combines a GPS receiver, a strapdown INS, and a cluster of CCD cameras mounted in a van. Requirements for the system are to provide precise positions at the 0.3 m level (RMS) of the cluster of CCD cameras. The digital imagery is then used to acquire positions of objects in the road corridor within a 50 m radius. These points are required with a 0.1 m relative accuracy. Tests have been conducted to confirm that the system can deliver these levels of accuracy, and further details are given in Schwarz et al. (1993). Another system which has been developed at Ohio State University is described in Bossler and Novak (1993).

5. MARINE APPLICATIONS

5.1 Water Level Profiling

Precise knowledge of water levels is essential for tidal studies and other hydrographic purposes such as the establishment of chart datums. GPS offers the possibility of determining water level profiles with a cm-level accuracy using carrier phase measurements on-the-fly. Accurate Bench Marks (B.M.'s) can also be established along the shores if an accurate geoid model is available. This is an important development since levelling operations may be prohibitive in remote areas such as the McKenzie River in Northern Canada. A feasibility study was conducted by The University of Calgary in conjunction with the Canadian Hydrographic Service (CHS) to obtain water level profiles and establish B.M.'s along an 80 km segment of the Fraser River, British Columbia. A first order levelling line along the river was used to assess the accuracy of the GPS-derived orthometric heights.

Results from the test are given in Lachapelle et al. (1993) and show that an accuracy of about 5.5 cm (RMS) could be achieved between the GPS-derived orthometric heights and the levelled B.M. height. Part of the error budget is due to the height transfer method (accurate to about 2 cm) as well as the geoid. Overall, the feasibility study was successful and the GPS-derived water level profiles are being analysed to determine the extent of tidal and other effects.

5.2 Attitude Determination

The use of attitude determination systems in the marine environment is important for the correction of hydrographic data from multi-beam acoustic surveying, for example. Traditionally, INS have been used since no other system has shown the capability of providing the level of required accuracy. However, INS drift over time so a cost-effective alternative is a GPS attitude determination system. These systems can either be self-contained units which is one receiver with several channels dedicated to each antenna (usually up to four), or may be comprised of several independent receiver systems.

The latter case was tested in June, 1993 in a joint effort between The University of Calgary and CHS. Four NovAtel GPSCard™ receivers were mounted on a hydrographic survey vessel with separations of 12 to 42 metres. Data was collected at a 10 Hz rate and post-processed to derive the ship's roll, pitch and azimuth. In the test, the ship's roll reached up to 10 degrees and when the GPS results were compared to the roll derived from an onboard INS, the RMS agreement was at the level of 2.7 arcminutes, with an agreement of 1.1 arcminutes during straight portions of the trajectory. Sample roll results are given in Figure 7. These results clearly show the feasibility of using multi-antenna GPS for accurate attitude determination. It is expected that these systems will be further utilized in the marine environment with the continual decrease in GPS receiver costs. See Lachapelle et al. (1994) for more information on the above test.

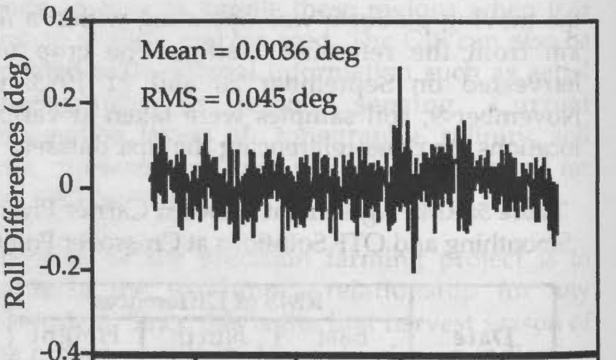


Fig.7: GPS-INS Roll Differences - June 93 Test

6. AIRBORNE APPLICATIONS

Many of the current airborne GPS applications centre around the photogrammetric and remote sensing areas as well as for precision landing of civilian aircraft. The latter application will not be discussed in this paper, but can be found in existing literature.

The University of Calgary has been involved with the development of GPS and GPS/INS systems for precise aircraft positioning and attitude determination aimed at the minimization or elimination of ground control in mapping applications. The accuracy requirements for such a system are generally dependent on the system in use as well as on the flying altitude. In general, the requirements are at the level of 10 cm or better for large-scale photogrammetry while they are relaxed to about 0.5 m for many remote sensing applications. GPS attitude is in general not sufficiently accurate for mapping applications, where the requirement may be in the order to 10-20 arcseconds. A description of the approach to GPS/INS integration for remote sensing can be found in Schwarz et al. (1993) and Sun et al. (1994).

Several flight tests have been carried out to confirm the feasibility of using GPS or GPS/INS for precise aircraft positioning. One of the difficulties is confirming the achievable accuracy during such tests since an independent system is required. One alternative is the use of inverse photogrammetry, where ground control is used to determine the coordinates of the GPS antenna which then can be compared to the estimated GPS coordinates, for example. Another concept is to build redundancy in the number of airborne and ground GPS receivers so independent aircraft trajectories can be compared. This concept is described in Shi and Cannon (1994) for the case when four receivers are used, two in the aircraft and two on the ground. In this test, baselines of over 100 km generated positioning accuracies at the 10-20 cm level when dual frequency carrier phase measurements are used with precise ephemerides.

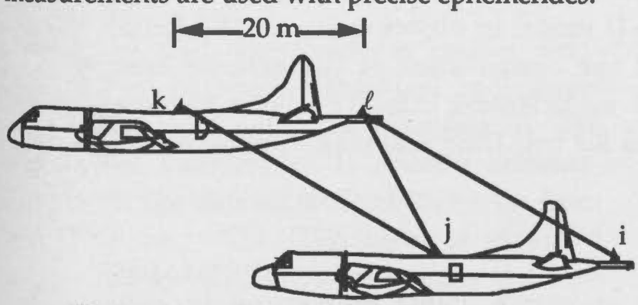


Fig.8: Aircraft-to-Aircraft Positioning Concept with Four Receivers

Many other applications in the airborne environment exist, with one being the precise positioning of one aircraft with respect to another. This work is being conducted between The University of Calgary and the U.S. for application in their magnetics program (Lachapelle et al., 1994). The concept is shown in Figure 8. Several tests have been conducted with the system and results show that cm-level positioning can be achieved between the two aircraft if an on-the-fly algorithm is used in the data reduction.

7. CONCLUSIONS

This paper summarized the current status of GPS and outlined the various classes of receivers that are currently available. With the arrival of the operational phase of GPS, and the reduction of GPS receiver costs, the number and breadth of GPS applications is growing at a fast rate. Various modes of acquiring data with GPS were reviewed. Several applications using GPS were presented with the main focus being on kinematic applications for land, marine and airborne environments. Overall, the pace of GPS activities is accelerating and it is expected that it will play an important continued role in the development of GIS.

REFERENCES

Alsip, D. (1993). Implementation of the U.S. Coast Guard's Differential GPS Navigation Service. Proceedings of the ION 49th Annual Meeting. Cambridge, June 21-23, pp. 707-716.

Bossler, J.D., K. Novak (1993). Mobile Mapping Systems: New Tools for the Fast Collection of GIS Information. Proceedings of GIS'93. Ottawa, March 23-25, pp. 306-315.

Cannon, M.E. (1990), High-Accuracy GPS Semikinematic Positioning: Modeling and Results, Navigation, Journal of the Institute of Navigation, Vol. 37, No. 1, pp. 53-64.

Cannon, M.E., and G. Lachapelle (1992), Analysis of a High Performance C/A Code GPS Receiver in Kinematic Mode, Navigation, Journal of the U.S. Institute of Navigation, Vol. 39, No.3, pp 285-299.

Cannon. M.E., E. Berry and M. King (1993), Testing a Lightweight GPS/GIS Receiver for Sub-Metre Accuracies, Proceedings of ION GPS-93, Salt Lake City, September 22-24.

Frodge, S.L., B. Remondi and D. Lapucha (1994), Results of Real-Time Testing and Demonstration of the U.S. army Corps of Engineers Real-time On-the-Fly Positioning System, Proceedings of the FIG XX Congress, Melbourne, March 6-12.

Gehue, H., G. Lachapelle, M.E. Cannon, T.W. Goddard and D.C. Penney (1994). GPS System Integration and Field Approaches in Precision Farming. Proceedings of the ION National Technical Meeting, San Diego, January 24-26 (in press).

Lachapelle, G. (1994). Marine Navigation and Surveying Developments. Proceedings of the ION National Technical Meeting, San Diego, January 24-26 (in press).

Lachapelle, G. R. Klukas, W. Qui and T.E. Melgaard (1994). Single Point Satellite Navigation Accuracy - What the Future May Bring. Proceedings of the IEEE PLANS-94, Las Vegas, April 11-15 (in press).

Lachapelle, G., C. Liu, G. Lu, Q. Weigen and R. Hare (1994). Water Level Profiling with GPS. Proceedings of the ION GPS-93, Salt Lake City, September 22-24.

Lachapelle, G., G. Lu and M.E. Cannon (1994), Attitude Determination with a Multi-Antenna GPS System for Hydrographic Applications, Contract report prepared for the Canadian Hydrographic Service (Atlantic Region).

Lachapelle, G., H. Sun, M.E. Cannon and G. Lu (1994), Precise Aircraft-to-Aircraft Positioning Using a Multiple Receiver Configuration, Proceedings of the ION NTM, San Diego, January 24-26, (in press).

McDonald, K. (1989), An Analysis of GPS Receiver Performance and Trends, Proceedings of the CPA Colloquium V, Calgary, October 4-6.

McLellan, J.F. and J.P. Battie (1994), OEM GPS Sensor Testing for Land and Marine Applications, Proceedings of the ION NTM, San Diego, January 24-26 (in press).

Melgaard, T., G. Lachapelle and H. Gehue (1994). GPS Signal Availability in an Urban Area - Receiver Performance Analysis. Proceedings of the IEEE PLANS-94, Las Vegas, April 11-15 (in press).

Schwarz, K.P., D. Lapucha, M.E. Cannon and H. Martell (1990), The Use of GPS/INS in a Highway Inventory System, Proceedings of the Fédération Internationale des Géomètre XIX International Meeting, Helsinki, Finland, Vol 5, pp. 237-249.

Schwarz, K.P., H.E. Martell, N. El-Sheimy, R. Li, M.A. Chapman and D. Cosandier (1993). VISAT- A Mobile Highway Survey System of High Accuracy. Proceedings of the IEEE Vehicle Navigation and Information Systems Conference, October 12-15, Ottawa, pp. 476-481.

Schwarz, K.P., M.A. Chapman, M.E. Cannon and P. Gong (1993), An Integrated INS/GPS Approach to the Georeferencing of Remotely Sensed Data, PE&RS, Vol. 59, No. 11, pp. 1667-1674.

Shi, J. and M.E. Cannon (1994). High Accuracy Airborne GPS Positioning: Testing, Data Processing and Results. Proceedings of the IEEE PLANS-94, Las Vegas, April 11-15 (in press).

Sun, H., M.E. Cannon, T. Owen and M. Meindl (1994), An Investigation of Airborne GPS/INS for High Accuracy Position and Velocity Determination, Proceedings of the ION NTM, San Diego, January 24-26 (in press).

Van Dierendonck, A.J. and P. Fenton (1992), Theory and Performance of Narrow Correlator Spacing in a GPS Receiver, Navigation, Vol. 39, No.3, pp.265-284.

L'UTILISATION DU SPG POUR LE GÉORÉFÉRENCEMENT PAR SIG : EXPOSÉ DE LA SITUATION ET APPLICATIONS

RÉSUMÉ

L'utilisation du SPG pour le géoréférencement vient d'amorcer une nouvelle ère grâce à la capacité opérationnelle du SPG et aux perfectionnements de la technologie des récepteurs et des stratégies de traitement du SPG. La situation actuelle des méthodes et des précisions SPG est exposée dans le contexte des applications SIG et les tendances en matière de matériel SPG sont aussi présentées. Plusieurs applications du SPG pour l'acquisition de données reliées spatialement sont présentées. Elles touchent, entre autres, à la cartographie de la salinité et à l'épandage d'engrais dans le domaine de l'agriculture. D'autres applications dans les milieux boisés, là où les zones d'ombre peuvent limiter la précision et l'applicabilité du SPG, sont également présentées. Le présent document termine en traitant de la tendance prévue en matière de traitement du SPG, en particulier du positionnement haute précision en temps réel et de ses répercussions dans le domaine du SPG.

THE EXTERIOR ORIENTATION OF DIGITAL IMAGES BY ROAD MATCHING

FAYEZ SHAHIN & KURT NOVAK

Dept. of Geodetic Science & Surveying
The Ohio State University
Columbus, OH 43210, USA
Commission II, Working Group 1.

KEY WORDS: B-splines, linear feature extraction, relational matching, automatic orientation.

ABSTRACT

The ultimate goal of digital photogrammetry is to produce maps automatically from digital images. For that purpose, the exterior orientation of the photos is required. It establishes the correspondence between the image and ground coordinates. Most approaches only address the relative orientation of the images. This paper presents a general procedure for the automation of the exterior orientation, which can be used for real time mapping applications. This procedure takes advantage of a mobile mapping system (GPSVan) developed and implemented at The Ohio State University's Center for Mapping. Various sensors collect information about the environment of highways from a moving van. A satellite (GPS) receiver mounted on the van determines the road alignment in the ground coordinate system. These alignments are mathematically represented by cubic B-splines to serve as a 3-D model of the roads on the ground. Digital images covering the same area are processed to find roads automatically. The extracted roads are also represented by cubic B-splines to serve as a model of the roads in image space. By applying relational matching and tree search methods, the best match between the roads in the digital images and their corresponding 3-D model in object space can be found. Thus, the correspondence problem can be solved and the computation of the exterior orientataion is possible. This approach is extremely efficient for orienting satellite images and small scale aerial photos. Therefore, it has a great potential for real time mapping applications.

1. INTRODUCTION

In recent years it has been shown that several important photogrammetric tasks, like the relative orientation of images, the aerial triangulation, and the derivation of digital elevation models, can be automated with very little or no human operator intervention. The main concern of solving these tasks automatically is to establish correspondence between conjugate points or features in overlapping images. Several algorithms and techniques, such as area based matching and feature based matching, have been used successfully to solve the correspondence problem between overlapping images [Schenk et al. 1991, Ackerman 1991].

Another group of tasks, including the exterior orientation of images, the geo-registration of images, and the mapping from orthoimages, have not reached advanced levels of automation. Operator intervention is still necessary in these tasks. This group of tasks has to be solved by determining the correspondence between an image and a model of the object space; for example, for the exterior orientation of an aerial photograph one has to determine a match between ground control points and their conjugate locations in the image. Finding such a match is a very difficult problem, because it requires matching the 3-D ground representation of control points, which is in vector form, to the 2-D image representation of the same points, which is in raster form.

In photogrammetric practice, the correspondence between image and object spaces is established by identifying and measuring control points in images manually. In many cases, such as in satellite images or small scale aerial photographs, it is very difficult to identify control points due to the low resolution of the images. This paper presents a general procedure for exterior orientation of images by matching features in the image with those on the ground. This approach is extremely efficient for orienting satellite images and small scale aerial photos. Thus, it has a great potential for real time mapping applications.

During the past three years a mobile mapping system has been developed and implemented by the Center for Mapping at The Ohio State University. It collects information about the environment of highways from a driving van. The GPS receiver mounted on the vehicle (GPSVan) determines the road alignment in a ground coordinate system [Novak 1991]. These road alignments provide a model of the road in object space. Digital images of the same area can be analyzed to extract roads automatically. By comparing roads on the ground and in the images, a correspondence between linear features in object and image spaces can be established automatically. Consequently, the automatic exterior orientation of digital images using linear features as control can be solved [Tankovich 1991, Mikhail 1993].

The system implemented for solving the correspondence problem between linear features in images and on the ground consists of three modules. These modules are the ground module, the image module, and the matching module. They are discussed in the following sections. Practical results are presented in section 5 of this paper.

2. THE GROUND MODULE

The GPSVan captures a huge number of discrete 3-D ground coordinates representing the centerline of a road. The goal of this ground module is to segment road centerlines and describe them mathematically. The first step is the detection of the critical (important) points of the road alignment. Next, the coefficients of the B-splines are computed such that the critical points serve as break points of the piecewise polynomials.

2.1 Detection of Critical Points

The detection of critical points for B-spline representation is a crucial step, because these points are representative of the original data. Redundant points should be deleted, and important features of the original data should be maintained.

The method implemented by the authors is based on the local properties of the curve. It depends on calculating the angles between two line segments $P_{i-1}P_i$ and P_iP_{i+1}. All points P_i for which this angle is less than a threshold are deleted. This method has the advantage of keeping only a few points for a straight line, and many points for lines of high curvature [Ballard and Brown 1982].

2.2 Cubic B-splines Representation

Splines are named after the draftman's device for drawing fair curves between specific points. The splines considered here are the well-known B- splines [Barnhill and Riesenfeld 1974]. Piecewise cubic polynomials in parametric form $C_i(t) = (x(t), y(t), z(t))$ define a curve as a function of the parameter t such that

$$C_i(t) = \sum_{i=0}^{4} V_i B_i(t)$$

where $V_i = (x_i, y_i, z_i)$ are the coefficients of the polynomial and $B_i(t)$ are the basis functions. For the curve to appear continiously smooth, it must have positional, first derivative, and second derivative continuity at the break points, also known as knots. The coefficients of the piecewise polynomials (B-splines) physically define the vertices of a polygon that guides the splines to trace a smooth curve (guiding polygon). These vertices are also called control points (see Fig. 1). The control points of B-splines are invariant under affine, and projective transformations. In addition, the errors incurred for a linear feature when assuming the invariance property of the B-spline under a perspective transformation are small, particularly in the case of small scale images [Cohen and Wang, 1994].

Cubic polynomials are most frequently used for splines since they are the lowest order in which curvature can change sign. Using the knot points P_i, which can be derived using the critical point detection algorithm mentioned earlier, the piecewise polynomial coefficients, which are also the polygon vertices V_i, are obtained by solving a set of simultaneous equations

$$\frac{V_{i-1}}{6} + \frac{2V_i}{3} + \frac{V_{i+1}}{6} = P_i \qquad 1 \leq i \leq n.$$

Since there are two fewer equations than unknowns, two conditions are added for the open curve case to ensure that the curvature is zero at both ends of the curve. By adding these conditions, the number of equations is equal to the number of unknowns, and the equations can be solved.

The iterative algorithm of Yamaguchi [Yamaguchi 1988] is used to compute the vertices of the guiding polygon. This algorithm is based on the idea that the set of equations satisfy the convergence condition of the Gauss-Seidel method, and that there is a special relationship among the unknowns (the coefficient matrix is circulant).

3. THE IMAGE MODULE

The goal of this module is to extract road segments automatically from digital images and to represent these segments by cubic B-splines to serve as a model of the roads in image space. Usually a road segment in an image contrasts sufficiently with its background, has a uniform width, and has sufficient length. The Duda road operator is employed to extract pixels that most likely belong to road segments. This low level step is followed by thinning and road tracking to group connected pixels as road segments. Finally, segments that belong together are grouped, and the gaps between them are filled meaningfully.

3.1 The Duda Road Operator

The Duda road operator (DRO) is used to make explicit the locations which are assigned the highest likelihood of road presence in the image. DRO masks are applied in four principal directions: horizontal, vertical, right diagonal, and left diagonal. For more details about DRO the reader is referred to [Duda 1973].

Once the image is convolved by the four masks, every nonborder pixel has four evaluation scores associated with it. The maximum score for every pixel is retained, and the final scores are thresholded to keep only the best, which are then interpreted as producing positive responses. This procedure is followed by thinning and road tracking. The road tracking aims at connecting pixels to road segments. Finally, each road segment is stored as a list of pixel coordinates; a size threshold is used to eliminate short segments.

3.2 Segment Grouping

The roads extracted by the previous operations are highly fragmented due to the effect of noise, poor contrast, and occlusions. Thus, there is a need for a procedure to group segments that belong together and fill the gaps between them in a meaningful way. The segment integration process is based upon two main features: proximity and alignment of pairs of segments. As a first requirement, segment l is considered to be a neighbor of segment q, if l and q are reasonably close to each other, as gauged by the distance between their nearest end points. The second requirement is that the trend or alignment of segment l (relative to the end that is nearest q) does not deviate too much from the alignment of segment q (at q's end nearest to l) [Vasudevan 1988]. After the valid neighbors of every segment are determined, they are grouped together as continuous road segments. Finally, the gaps between the grouped segments are filled by the process of fitting cubic B-splines.

4. THE MATCHING MODULE

This module aims at finding the best matches between the roads extracted from digital images and their corresponding 3-D models in object space. The roads in both image and object spaces are represented by parametric cubic B-splines. These splines have the property of shape invariance under projective transformation. Thus, the coefficients of these splines form the primitives for the matching process. The search for the best match is conducted using tree search methods.

4.1 Primitives and Relations

The coefficients (vertices) of the splines form the primitives, while the distances between the vertices describe the interrelationships between these primitives. The segmented roads in object space are represented by a sequence of vertices of the guiding polygon as follows: a vertex O_i is represented as $(X_i, Y_i, Z_i, \theta_i)$, where X_i, Y_i, Z_i are the coordinates of the vertex and θ_i is the angle it encloses. The segmented roads in image space are also represented by a sequence of the vertices of the guiding polygon: a vertex L_i is represented as (x_i, y_i, α_i), where x_i, y_i are the image coordinates of the vertex and α_i is the angle it encloses. The distance T_i between two successive vertices in object space describes the relationship between these vertices. In addition, the distance R_i between two successive vertices in image space describes the relationship between these vertices. The distances in image and object space are related through the image scale.

4.2 The Matching Problem

The matching problem between the object model $M_o(O,T)$, where O and T are the vectors of primitives and relationships of the object model, and the image model $M_i(L,R)$, where L and R are the vectors of primitives and relationships of the image model, is simply to find a mapping f between primitives O and L. Since there may be many possible mappings $f : L \rightarrow O$, a measure has to be introduced, which evaluates the quality of the mapping between two primitives. Intuitively, the evaluation of a mapping should depend on the similarity of the attribute values of the corresponding primitives and relations. The approach to find the best mapping is called the inexact cosistent-labeling problem [Shapiro and Haralick, 1985].

Inexact consistent-labeling utilizes the concept of a cost function. For every possible mapping between the object model and the image model, a cost is considered based on the similarity of the corresponding primitives and relations. The best mapping is the one that minimizes the cost function. It can be found using backtracking search with forward checking [Haralick and Elliott, 1980].

4.3 Backtracking with Forward Checking

The backtracking tree search begins with the first unit of L. This unit can potentially match many labels in O. Each of these potential assignments is a node at level 1 of the tree. The algorithm then starts to construct the children of the first node, which are nodes that map the second unit of L to each possible label of O. A cost is associated with each label. The paths from the root node to any successful node are the consistent labellings. The best mapping is the path which minimizes the cost. Forward checking is used to cut down the search time by reducing the size of the tree that is searched. This checking is based on the idea that once a unit-label pair is selected at a node in the tree, the constraints imposed by the relations cause the selection of some future unit labels to become impossible.

5. EXPERIMENTS & RESULTS

Fig. 2 shows a digitized aerial image, the extracted linear features (DRO Image), the feature image model, and the road ground model. Fig. 3 shows a SPOT image, the extracted linear features (DRO Image), the feature image model, and the road ground model. Some roads were not extracted by the Duda road operator (DRO) due to the lack of contrast. In addition, other linear features and artifacts in the images were extracted. The linear features with sufficient length were matched to the road ground models. Hence, only roads that are represented in the ground model will have an acceptable match in the image. An acceptable match is the one whose cost function is minimum.

The cost function depends on the similarity between the attributes of the primitives, the similarity between the relations, and the number of missing vertices (wild cards). A larger weight is given to the attributes of the primitives, since they describe the general characterstices of the linear feature. Furthermore, the number of wild cards should not exceed a certain number. By adopting this methodology, the matched roads for both the aerial and SPOT images were found and are shown in Fig. 2 and Fig. 3. The matched roads were investigated manually and it was found that they correspond to their conjugate roads in the ground model. The conjugate roads in the image and on the ground can be used as linear control to perform image resection (Exterior Orientation).

6. CONCLUSION

The implemented methodology is unique in its attempt to automate the exterior orientation of digital photos. It has both practical and scientific significance. Very few in the photogrammetric community tried to solve the exterior orientation automatically. Most previous research has focused on the relative orientation of stereo-pairs or blocks of photographs. Furthermore, for orienting satellite images or small scale aerial images, existing maps (USGS quads) are often used to find control features on the ground. The accuracy of these maps is roughly 15m, which is not sufficient for most mapping applications, while the accuracy of the roads captured by the GPSVan is about 2m. By using the implemented methodology, more accurate ground control can be obtained, as compared to the existing manual methods. Another major problem with using digitized maps for ground control is that they were generalized and many features existing in the images do not appear on the map or are offset on purpose. Therefore, it makes sense to drive the van, to create an accurate road map dynamically, and to use it for the orientation of the images, as well. We believe that this approach has a great potential for real time mapping applications.

References

[1] Ackerman, F. and P. Krzystek [1991]: MATCH-T: Automatic Mensuration of Digital Elevation Models. Presented paper to the 3rd Technical Seminar of the Sociedad Espanola de Cartografia Fotogametria y Teledeteccion, April 2 1991.

[2] Ballard, D. H. and C. M. Brown [1982]: *Computer Vision.* Prentice Hall, Englewood Cliffs, NJ, 1982.

[3] Barnhill, R. E. and R. F. Riesenfeld [1974]: *Computer Aided Geometric Design.* Academic Press, New York, 1974.

[4] Cohen, F. S. and J. Wang [1994]: Part I: Modeling Image Curves Using Invariant 3-D Object Curve Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 16, pp. 1-12, 1994.

[5] Duda, R. O. and P. E. Hart [1973]: *Pattern Classification and Scene Analysis.* Wiley, New York, 1973.

[6] Haralick, R. M. and G. L. Elliott [1980]: Increasing Tree Search Efficiency for Constraint Satisfaction Problems. *Artificial Intelligence,* vol. 14, pp. 263-313, 1980.

[7] Mikhail, E. M. [1993]: Linear Features for Photogrammetric Restitution and Object Completion. In: *Proceedings of SPIE,* vol. 1944, pp. 16-30, 1993.

[8] McKeown, D. M. and J. F. Pane [1985]: Alignment and Connection of Fragmented Linear Features in aerial imagery. Dept. of Computer Science, Carnegie-Mellon University, Tech. Rep. CMU-CS-85-122, 1985.

[9] Novak, K. [1991]: The Ohio State University Highway Mapping System: The Positioning Component. In: *Proceedings of the Institute of Navigation Conference,* Williamburg, VA., 1991.

[10] Shapiro, L. G. and R. M. Haralick [1985]: A Metric for Comparing Relational Structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 7, pp. 90-94, 1981.

[11] Schenk, T., J-C Li and C. Toth [1991]: Towards an Autonomous System for Orienting Digital Stereopairs. *Photogrammetric Engineering & Remote Sensing,* vol. 57, pp. 1057-1064, 1991.

[12] Tankovich, J. [1991]: *Multiple Photograph Resection and Intersection Using Linear Features.* M.S. Thesis. Dept. of Geodetic Science and Surveying. The Ohio State University, Columbus, Ohio, 1991.

[13] Vasudevan, S., R. L. Cannon, J.C. Bzedek and W.L. Cameron [1988]: Heuristic for Intermediate Level Road Finding Algorithms. *Computer Vision, Graphics, and Image Processing,* vol. 44, pp. 175-190, 1988.

[14] Vosselman, G. [1992]: *Relational Matching.* Springer-Verlag, New York, 1992.

[15] Yamaguchi, F. [1988]: *Curves and Surfaces in Computer Aided Geometric Design.* Springer-Verlag, New York, 1988.

LIST OF FIGURES

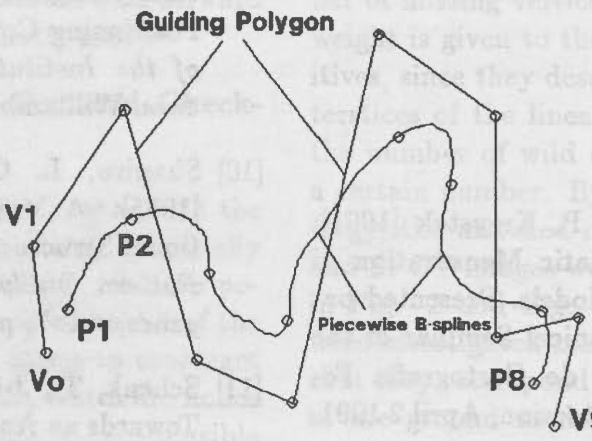


Fig. 1: Control Points, Knots of B-splines.

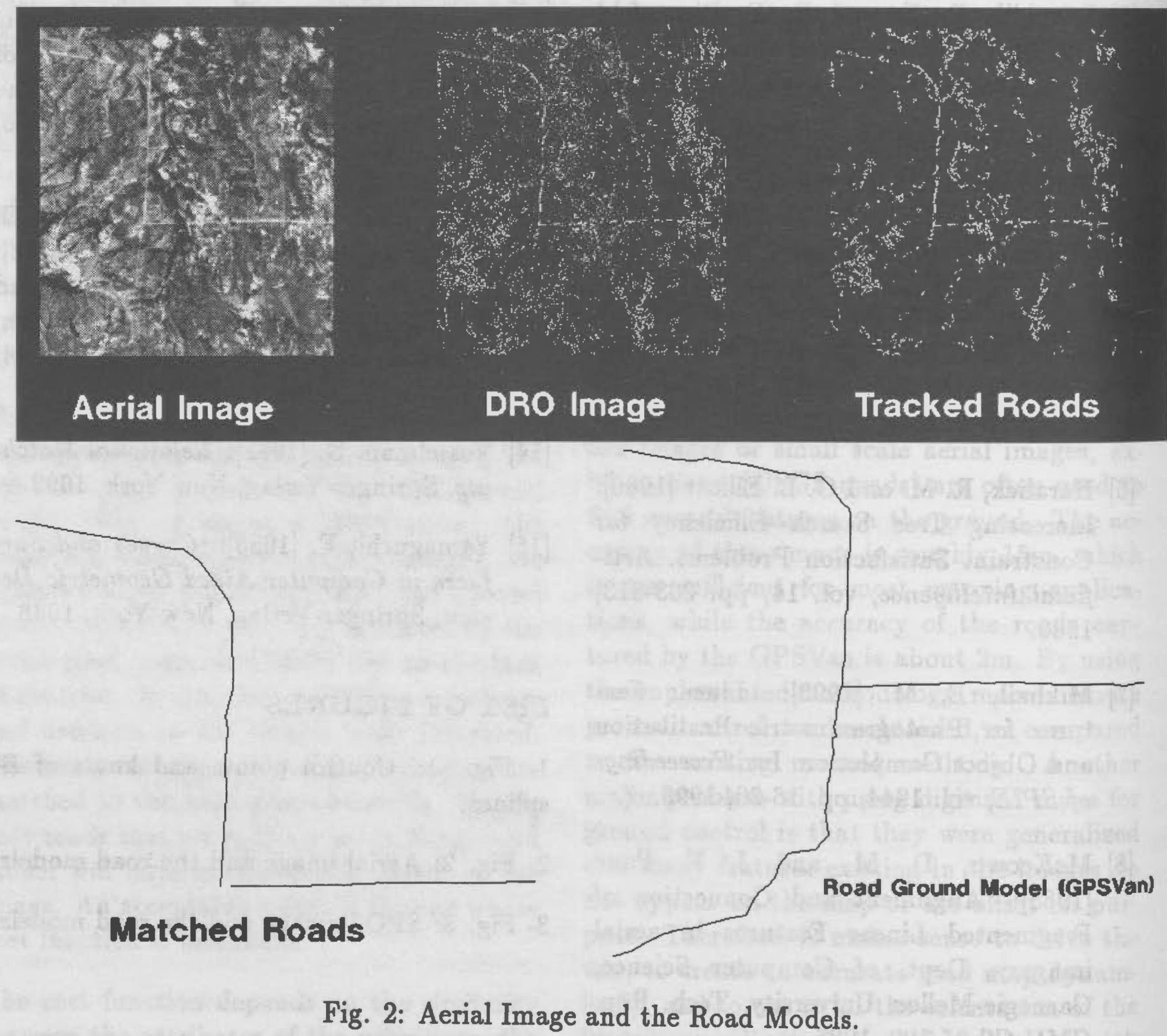


Fig. 2: Aerial Image and the Road Models.

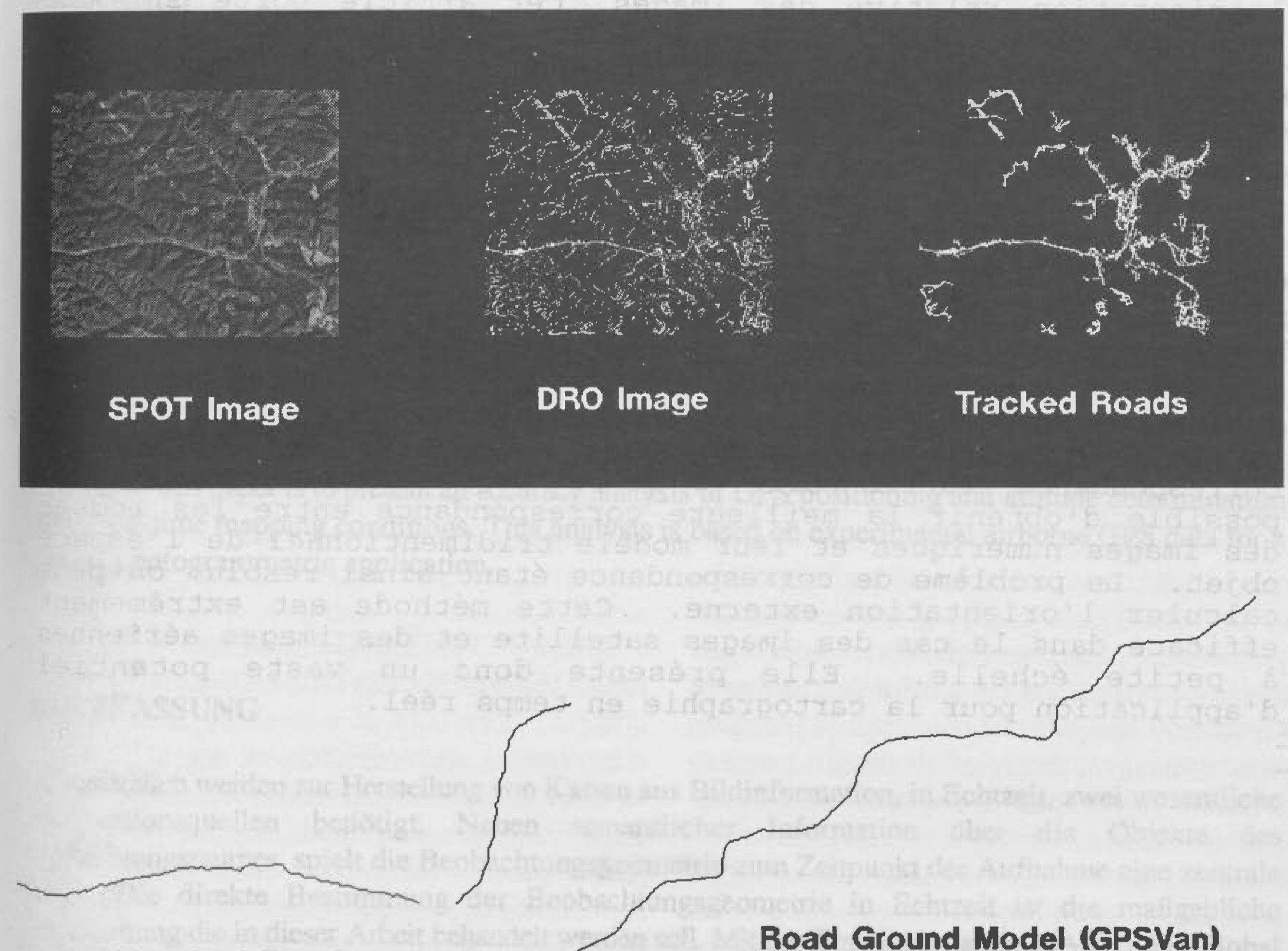


Fig. 3: SPOT Image and the Road Models.

Orientation externe des images numériques par correspondance des routes

RÉSUMÉ

Le but ultime de la photogrammétrie numérique est de produire automatiquement des cartes à partir d'images numériques. Pour atteindre cet objectif, l'orientation externe des photographies est nécessaire car elle établit la correspondance entre l'image et les coordonnées-terrain. En photogrammétrie, on obtient cette correspondance en identifiant et mesurant manuellement les points de contrôle. On a très rarement tenté jusqu'ici de mettre au point une méthode visant à automatiser le processus d'orientation des images numériques. La plupart des méthodes ne concernent que l'orientation relative des images. Cet article porte sur une technique générale d'établissement automatique de l'orientation externe, applicable à la cartographie en temps réel. On se sert d'un système de cartographie mobile (GPSVan) conçu et mis au point par le Ohio State University's Center for Mapping. Ce système recueille des informations sur l'environnement des routes à partir d'un fourgon. Un récepteur (SPG) satellite installé sur le véhicule détermine le tracé des routes dans le système de coordonnées-terrain. Ces tracés sont représentés sous forme mathématique par des B-splines cubiques qui permettent d'obtenir un modèle en trois dimensions des routes au sol. Des images numériques couvrant la même zone sont traitées afin de trouver les routes de manière automatique. Les routes extraites sont également représentées par des B-splines cubiques, qui forment un modèle de route dans l'espace image. À l'aide des méthodes de correspondance relationnelle et de recherche arborescente, il est possible d'obtenir la meilleure correspondance entre les routes des images numériques et leur modèle tridimentionnel de l'espace objet. Le problème de correspondance étant ainsi résolu, on peut calculer l'orientation externe. Cette méthode est extrêmement efficace dans le cas des images satellite et des images aériennes à petite échelle. Elle présente donc un vaste potentiel d'application pour la cartographie en temps réel.

EXTERIOR ORIENTATION FOR AIRBORNE REAL TIME MAPPING

HOLGER SCHADE
Institute for Photogrammetry
University of Stuttgart
Keplerstr. 11
D-70174 Stuttgart
GERMANY

KEY WORDS: Exterior Orientation, GPS, GPS Attitude, Real-Time Mapping, Accuracy Analysis

ABSTRACT

Real-time mapping applications from aerial sensor systems, require in principle two major data sources. On one hand it is important to gather knowledge about the observation geometry during the flight and on the other hand it is necessary to acquire as much information as possible about the object space which is to be mapped.

This paper will concentrate on the acquisition of information about the observation geometry. With the advent of the Global Positioning System (GPS) the real-time determination of an aircrafts position with high accuracy (< 1m) has become available. Besides positioning, new developments of multi-antennae GPS receiver opened up the opportunity to measure also the aircrafts attitude, with an accuracy of less than 0.5 degrees in real-time. Hence, with GPS it has become possible to determine the entire exterior orientation of a moving aircraft in real time with high precision. The concern of this paper is to present an accuracy analysis of GPS positioning and attitude determination under real-time mapping conditions. This analysis is based on experimental airborne GPS data for a typical photogrammetric application.

KURZFASSUNG

Grundsätzlich werden zur Herstellung von Karten aus Bildinformation, in Echtzeit, zwei wesentliche Informationsquellen benötigt. Neben semantischer Information über die Objekte des Beobachtungsraumes, spielt die Beobachtungsgeometrie zum Zeitpunkt der Aufnahme eine zentrale Rolle. Die direkte Bestimmung der Beobachtungsgeometrie in Echtzeit ist die maßgebliche Fragestellung die in dieser Arbeit behandelt werden soll. Mit der Entwicklung des NAVSTAR/Global Positioning System bietet sich nun erstmals die Möglichkeit die Parameter der äußeren Orientierung direkt mit hoher Genauigkeit (< 1m) zu bestimmen. Neben der Positionierung, kann die Neigung eines Flugzeuges durch die Nutzung von Multi-Antennen GPS Empfängern ebenfalls direkt in Echtzeit mit Genauigkeiten im Bereich von einigen Zehntel Gon bestimmt werden. In dieser Arbeit wird eine Genauigkeitsanalyse der GPS Positionierung und Neigungsbestimmung unter den Rahmenbedingung einer Echtzeit-Kartierung diskutiert. Die Analyse basiert auf experimentellen Flugdaten einer photogrammetrischen Befliegung.

1. INTRODUCTION

With the extended spreading of geographic information systems (GIS) high demands are made on photogrammetry and mapping. As the underlying data is of utmost importance for geographic information systems, the need for a fast and cost-effective data acquisition changes the pre-requirements and working procedures in photogrammetry and mapping tremendously. The new photogrammetric and mapping methologies are characterized by:

- kinematic methods
- sensor fusion
- increased automatization
- real-time applications

From all the above mentioned characteristics the real-time aspect is by far the most demanding, as it implicitly contains all the other mentioned features. In real-time photogrammetry and mapping, data acquisition is usually kinematic. Further, the real-time aspect can often only be solved if a full automatization of all data processing steps can be guaranteed. And this automatization can only be achieved if as much information as possible can be gathered from a multi-sensor system.

In principle, two major problems have to be solved in real time mapping applications. On one hand the parameters of exterior orientation have to be determined for the full description of the observation geometry (see fig. 1). And, on the other hand, as much information as possible about the object space is required for the real-time, automatic reconstruction and classification of relevant objects. The automatic reconstruction and classfication of objects from image or other data sources is a highly complex problem and for the time beeing it can only be solved in real-time for some specific applications. This automatic interpretation step is a highly complex research topic on its own and it will not be treated in the remainder of this paper. Rather, the paper will concentrate on the real time determination of exterior orientation parameters with the aid of the NAVSTAR/Global Positioning System (GPS). Again, the problem of determining the exterior orientation of a sensor platform can be divided in two seperate tasks. The exterior orientation of airborne sensor systems consists of a positioning part for the determination of the projection center coordinates (X_0, Y_0, Z_0) and an attitude part where we have to solve for the attitude (ω, ϕ, κ) of the sensor system with respect to a known coordinate system.

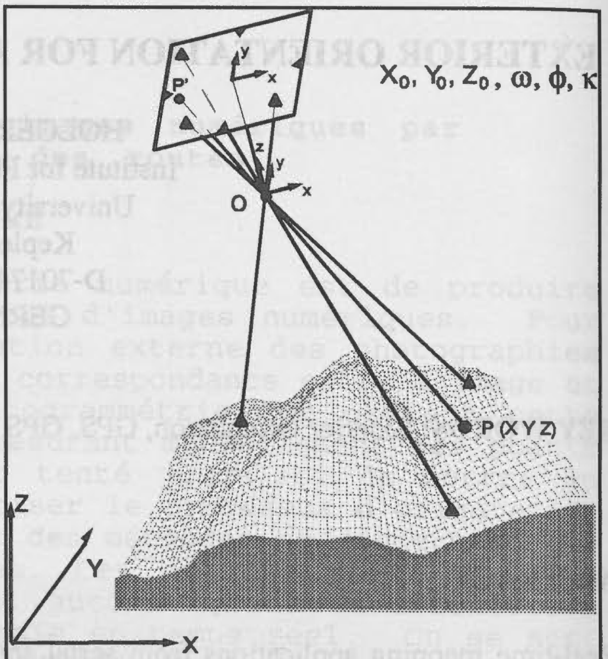


Figure 1 Exterior Orientation Parameters

2. REAL-TIME DETERMINATION OF PROJECTION CENTER COORDINATES

The precise determination of the projection center coordinates of airborne sensors has been significantly simplified by the recent development of the NAVSTAR/Global Positioning System (GPS). Although, GPS real-time, high precision positioning has become routinely available for airborne photogrammetric and mapping applications, extreme care has to be taken especially for large scale mapping. Depending on the map scale the positioning accuracy requirements for mapping applications can be fairly stringent (see table 1).

map scale	image scale	required σ_{GPS} X,Y [m]	required σ_{GPS} Z [m]
1:100000	1:100000	27	6,2
1:50000	1:65000	11	2,4
1:25000	1:40000	6,8	0,9
1:10000	1:25000	1,7	0,5
1:5000	1:12000	0,9	0,05

Table 1 Required Accuracies for different Map Scales

From table 1 it can be seen that the specifications for large scale mapping get more demanding as the map scale increases. The table has been mainly compiled for digital or analogue aerial cameras and it does not include the accuracy requirements for other airborne remote sensing sensors (e.g. CASI) or non-imaging sensors (eg. LASER). Nevertheless, the needed positioning accuracy will rarely be higher than the one for large scale (1:5000) mapping.

As the real-time determination of the projection center coordinates with GPS can be affected by large systematic errors (e.g. orbit, system time, atmospheric effects, multipath) the handling of the systematic error effects in the ranging model has to be of special concern. The most effective way of compensating these systematic errors is by using differential GPS (DGPS). For a detailed description of differential GPS applications please refer to BLACKWELL [1986]. The principle of DGPS is to observe the systematic error effects on a stationary receiver which is located over a known point and transmitting correction values to the moving receiver. For real-time applications the transmission of the correction data has to be done with telemetry links. Conventionally, these telemetry links are in the HF-UHF range due to legal constraints, transmission speed and line of sight problems. The information which is transmitted via the telemetry links can vary significantly, and it can range from simple coordinate corrections, over range corrections up to the full observation set on the reference station. An effort to standardize the transmitted values has been done by KALAFUS ET AL. [1986]. Using this technique most error sources related to the satellite and the signal transmission can be reduced. The size of the reduction is mainly dependent on the reference station receiver - moving receiver seperation. Receiver dependent error sources may be cancelled by differencing observations between satellites. The differencing principle can be used on all GPS observation types. In the remainder of this paper we refer to pseudorange and carrier phase observations in the sense of double differenced (between-station, between-satellite) observations.

The GPS observation types which are available and suitable for real-time positioning purposes are pseudoranges on the C/A-Code frequency as well as carrier phase observations on the L1- and L2-frequencies. The P-Code pseudoranges can only be used for military users and should not be considered available for most of conventional mapping applications. Further a combination of the two signal types, pseudorange and carrier phase, can also be used for positioning. However, the suitability of the different signal types under real-time, airborne kinematic conditions has to be carefully analyzed, as the accuracy and the processing requirements for the observation types vary significantly.

The usage of pseudorange data for real-time positioning is straight forward, as the pseudorange signal has been directly designed for this purpose. Using the observations from four satellites and the mathematical relation from equation 1 the position of a moving vehicle can be determined by solving for the position unknowns (X,Y,Z) and the receiver clock error (dt). It has to be stressed that all parameters of equation 1 are double differenced as explained in the previous chapter.

$$\rho = s\ (X,Y,Z) + cdt + d_{atmo} + \epsilon \qquad (1)$$

The remaining formula descriptors are ρ which is the double differenced pseudorange observation, *c* the constant speed of light, d_{atmo} the remaining atmospheric signal delay after double differencing and ϵ which is the observation noise. *s* which implicitly contains the moving station coordinates, is the double difference range which describes the double differenced distance between the satellite-receiver combination. The noise of the pseudorange observations is the critical part in real-time mapping applications. The measurement accuracy which can be achieved with double-differenced C/A-Code observations is conventionally in the range of 3-5 m, resulting in a positioning accuracy of 5-10 m. Figure 2a shows the position differences of the projection center coordinates determined with GPS pseudorange observations and conventional aerial triangulation. From this figure it can be seen, that the positioning accuracy with C/A-Code pseudoranges is sufficient only for small scale mapping (< 1:50000). Recent developments to minimize the measurement noise, with narrow correlation techniques might improve the situation considerably. Observation accuracies for C/A-Code observations in the range of 5-30 centimeters have been reported by LACHAPELLE ET AL. [1992], resulting in positioning accuracies of a few decimeters. As soon as this technology is routinely available the simple data handling and data processing makes this observation type highly interesting for small and medium scale real-time mapping.

The stringent accuracy requirements for large scale mapping makes the use of carrier phase observations necessary. In the case of carrier phase observations the incoming sine wave is measured against a reference wave which is generated by the GPS receiver. The problem with this measurement principle, is that the correlation of the two signals is ambiguous, and only the fraction of the wavelength, by which the incoming signal is shifted, can be measured. Once the receiver locks onto the signal, the cycle counts are updated and a range difference with respect to the initial epoch can be measured. Equation 2 shows the observation equation for the carrier phase observation (λ beeing the wavelength of the carrier phase observations).

$$\rho = s\ (X,Y,Z) + cdt + d_{atmo} + \lambda N + \epsilon \qquad (2)$$

If the initial ambiguity N can be successfully determined the inherent measurement accuracy (<2 mm) of the phase observations is sufficient for all map and image scales, resulting in a sub-decimeter position accuracy. Hence, especially for large scale mapping the initial ambiguity N has to be determined to exploit the high positioning accuracy with carrier phase observations. Further, it has to be kept in mind that the ambiguity has to be reinitialized as soon as the receiver does not update the cycle count correctly (cycle slips) or if the receiver can not track the satellite signal continuously (loss of lock e.g. due to signal obstructions). The need to determine the initial ambiguity complicates the use of carrier phase observations for real-time applications extremely. In principle, the initial ambiguities can be estimated and fixed at the beginning of a continuous sequence of observations in a static initialization, but due to banking angles in flight turns and the highly kinematic environment, losses of phase lock and cycle slips are frequent in airborne, real-time mapping applications. Hence, in most cases there is a need to re-initialize the ambiguities while the aircraft is moving. In recent publications several authors proposed methods for a real-time initialization of the ambiguities ("ambiguity resolution on the fly"), based on statistical searching algorithms (e.g. HATCH [1990], FREI/BEUTLER [1990], SCHADE [1992]). The mentioned algorithms share some basic principles to distinguish between the correct cycle ambiguities and the incorrect ones. Usually, a n adjusted pseudorange position and its associated covariance information is used as a searching cube in which the potential solution

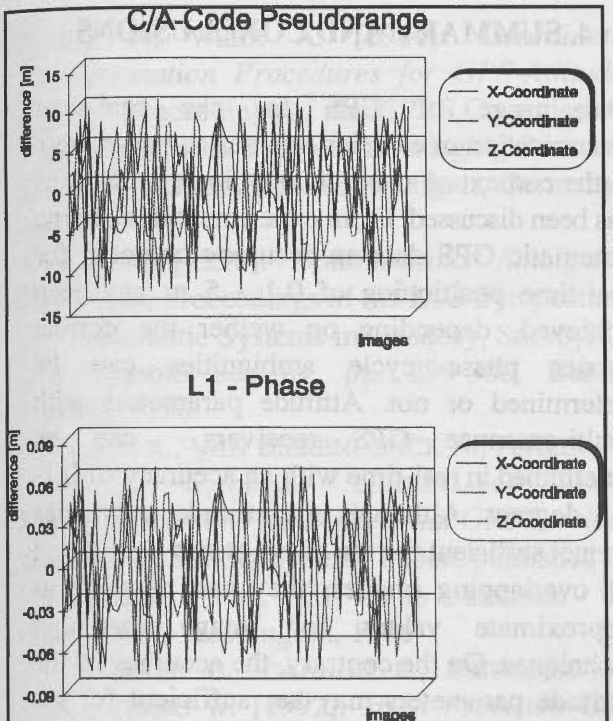


Fig. 2 Comparison between photogrammetrically and GPS determined projection center coordinates

has to be. Mostly, the variance of the carrier phase observation is used in statistical tests to determine wether a set of carrier phase ambiguities is potentially correct or false. Further, the computer processing requirements are very large, because often more than several hundredthousand possible ambiguity combinations have to be tested in real-time. Although, the ambiguity resolution on the fly has proven its applicability in numerous tests it is important to mention, that for the time being the reliability of these algorithms can not be guaranteed. Especially, if only single frequency C/A-Code receivers are used and the baselines between the moving and reference receivers are large, the convergence to incorrect ambiguity solutions are very likely. Nevertheless, if one is able to find the correct ambiguities, GPS carrier phase observations can provide sufficient positioning accuracy for all real-time mapping applications. Figure 2b shows the position differences of the projection center coordinates determined with GPS carrier phase observations and conventional aerial triangulation.

3. REAL-TIME DETERMINATION OF SENSOR ATTITUDE PARAMETERS

Apart from the determination of the projection center coordinates of sensors, the reconstruction of objects on the earth's surface from sensor information, requires the measurement of the attitude angles of the sensor with respect to a known coordinate system. For the direct measurement of real-time attitude parameters only a few sensors are available (e.g. INS). Recent developments of multi-antennae GPS receivers have added a further potential method for the real-time, kinematic attitude determination of sensor systems. Attitude determination with GPS is based on the interferometric measurements of GPS carrier phase data. The phase difference $\Delta\Phi$ which can be observed between two antennas results from the range difference between the satellite to the antennas. As the distance between an antenna and the satellite is rather large (> 20000 km) compared to the short distance between the antennas (< 20 m), the incoming phase signal can be assumed to be parallel. Therefore, the phase difference is just dependent on the baselength B and the angular position γ of the satellite with respect to the baseline between the antennas (see Eq. 3).

$$\gamma = \arccos \frac{\Delta\Phi * \lambda}{2\Pi * B} \qquad (3)$$

The phase difference has to be measured with the highest possible accuracy as small errors in the phase difference may result, depending on the baseline length, in large attitude errors (1cm/10m = 0.1°). The baseline B between the antennas can be measured prior to the mapping mission in a calibration measurement with conventional survey methods. It is clear that the above equation only holds if the carrier phase cycle ambiguity for the interferometric measurement has been determined correctly. However, here the resolution of the correct ambiguities is simplified compared to conventional positioning. In the attitude computation algorithms, the known baseline length between the antennas can be used as additional constraint for the ambiguity resolution. Due to these constraints the number of potential solutions which have to be tested is reduced by a significant amount and hence the on-board processing requirements is minimal. Further, the resolution of the ambiguities is much more reliable as for the positioning case. However, for the airborne, kinematic attitude determination the baseline can not enter the attitude computations as error-free. Due to structural deformations of the aircraft body and wings during the flight, caused by the aerodynamic behaviour, the baseline length can vary significantly (see also COHEN/PARKINSON [1992]).

It has been mentioned earlier, that the highest measurement accuracies are required for GPS attitude determination. Therefore, error effects like multipath, antenna phase center variations, wing flex and receiver noise have to be kept as small as possible. A more detailed review about the error effects and the data-processing techniques for real-time, kinematic attitude determination with GPS can be found e.g. in HARTL/WEHR [1986]. Due to its simple handling its robustness and the minimal data processing requirements GPS attitude determination is highly suitable for real-time mapping applications. The potential of this sensor has been analyzed under conventional photogrammetric conditions, yielding attitude accuracies in the range of 0.1 - 0.2 degrees. Figure 3 shows a comparison of GPS attitudes and the attitudes computed from a conventional photogrammetric block adjustment. A more detailed description of this testflight can be found in SCHADE ET AL. [1993].

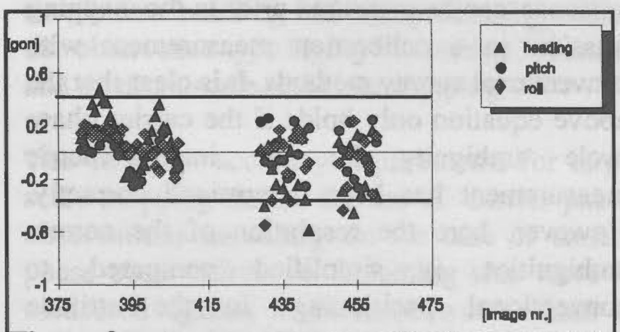


Figure 3 Differences between Photogrammetric and GPS Attitudes

4. SUMMARY AND CONCLUSIONS

The usage of GPS for the real-time determination of exterior orientation parameters in the context of airborne mapping applications has been discussed. From experimental, airborne kinematic GPS data an accuracy potential for real-time positioning of 0.1 - 5 m has been achieved, depending on wether the correct carrier phase cycle ambiguities can be determined or not. Attitude parameters with multi-antennae GPS receivers can be determined in real-time with an accuracy of 0.1-0.3 degrees. Although, the attitude accuracies are not sufficient for the direct stereo-evaluation of overlapping images, they can be used as approximate values for image matching techniques. On the contrary, the accuracy of the attitude parameters may be sufficient for the direct exterior orientation of non-imaging sensors, monoplotting purposes or the direct exterior orientation of airborne-remote sensing sensors. In conclusion, it has become obvious that with the advent of the Global Positioning System the real-time exterior orientation of airborne sensor systems can be achieved with sufficient accuracy for most mapping applications.

5. BIBLIOGRAPHY

BLACKWELL E.G. [1986]: *Overview of Differential GPS Methods*, Global Positioning System, Papers published in Navigation, Vol. III, The Institute of Navigation, Washington, 1986

COHEN C.E, PARKINSON B.W. [1992]: *Aircraft Applications of GPS Based Attitude Determination*, Proceedings of the ION GPS-92, Fifth International Technical Meeting of the Satellite Division of the Institute of Navigation, pp.775-782, Albuquerque,USA 1992

FREI E., BEUTLER G. [1990]: *Rapid Static Positioning Based on the Fast Ambiguity Resolution Approach "FARA": Theory and Results*, manuscripta geodaetica, Vol. 15, Springer Verlag, pp. 325-356, 1990

HARTL P., WEHR A. [1986]: *Coordinate Transformation Procedures for GPS-Attitude Control,* Proceedings of the ISPRS Commission I Symposium on Progress in Imaging Sensors, ESA SP-252, pp.215-226, Stuttgart, Germany, 1986

HATCH R. [1990]: *Instantaneous Ambiguity Resolution,* Proceedings of the IAG Symposium 107 Kinematic Systems in Geodesy, Surveying and Remote Sensing, pp. 299-308, Banff, Canada, 1990

KALAFUS R., VAN DIERENDONCK A., PEALER N. [1986]: *Special Committee 104 Recommendations for Differential GPS Service,* Global Positioning System, Papers published in Navigation, Vol. III, The Institute of Navigation, Washington, 1986

LACHAPELLE G., CANNON E., ERICKSON C., FALKENBERG W. [1992]: *High Precision C/A Code Technology for Rapid Static DGPS Surveys,* Proceedings of the 6th International Geodetic Symposium on Satellite Positioning, Columbus, Ohio, USA, 1992

SCHADE H. [1992]: *Reduction of Systematic Errors in GPS-Based Photogrammetry by Fast Ambiguity Resolution Techniques,* International Archives of Photogrammetry and Remote Sensing, Vol. XXIX, Part B1, pp. 223-228, 1992

SCHADE H., LACHAPELLE G., CANNON E. [1993]: *An Accuracy Analysis of Airborne Kinematic Attitude Determination with the NAVSTAR/Global Positioning System,* SPN-Zeitschrift für Satellitengestützte Positionierung, Navigation und Kommunikation, Vol. 3/93, pp. 90-95,1993

Orientement externe pour système aéroporté de cartographie en temps réel

RÉSUMÉ

Les applications de cartographie en temps réel nécessitent en principe deux sources principales de données. Il faut d'une part connaître les paramètres de géométrie d'observation pendant le vol et recueillir d'autre part le plus d'informations possible en ce qui a trait à l'espace objet à cartographier.

Le présent document traite de la saisie de l'information relative à la géométrie d'observation. Depuis l'avènement du système de positionnement global (GPS), il est devenu possible de déterminer en temps réel et avec une grande précision (< 1 m) la position d'un aéronef. La mise au point récente des récepteurs GPS multiantennes vient en plus compléter cette information en donnant aussi l'attitude de l'aéronef, avec une précision angulaire de l'ordre de 0,5 degrés et également en temps réel. Le système de positionnement global permet donc maintenant de déterminer en temps réel tout l'orientement externe d'un aéronef en mouvement. Le présent document se veut une analyse des problèmes de précision du système de positionnement global et de détermination de l'attitude de l'aéronef de cartographie, dans les conditions types d'un vol normal et en temps réel. Les données de cette analyse ont été recueillies lors de vols d'un appareil expérimental équipé de matériel photogrammétrique.

A PRECISE POSITIONING/ATTITUDE SYSTEM IN SUPPORT OF AIRBORNE REMOTE SENSING

K.P. Schwarz, M.A. Chapman, M.E. Cannon, P. Gong, D. Cosandier
Department of Geomatics Engineering, The University of Calgary
2500 University Drive NW, Calgary, AB, Canada T2N 1N4
Tel. (403) 220-7377 Fax: (403) 284-1980

ISPRS, Commission II, WGII/1 - Session 2

KEY WORDS: Precise position and attitude system, geo-referencing, INS/GPS integration, pushbroom scanner, CCD frame imager, SAR, calibration of airborne sensors

ABSTRACT

Research in airborne remote sensing at The University of Calgary has as one of its goals the development of a precise positioning and attitude system that can be used with a variety of airborne sensors and will ultimately eliminate the need for ground control for georeferencing. In this paper, accuracy requirements for such a system are discussed, different sensor configurations are described, and results of the U of C prototype development are analyzed. Applications can be subdivided into three major groups: those where precise positioning is the major requirement, such as photogrammetric applications; those where both position and attitude are required with high accuracy, such as pushbroom imaging applications; and those where accurate velocity estimation for motion compensation is also needed, as in radargrammetric applications. Sensor configurations for different applications will be discussed and first results of airborne tests will be briefly reviewed.

1. INTRODUCTION

Airborne remote sensing considerably extends the capabilities of satellite remote sensing in terms of resolution and operational planning. Where satellite remote sensing at best achieves accuracies of 10 -15 m, airborne remote sensing has the potential of achieving accuracies at the decimeter level in position. Whereas the usefulness of satellite remote sensing is often restricted by the images available for a certain area and the extent of the intervening cloud coverage, there are no such limitations in airborne remote sensing. It is, therefore, possible to optimize the required result by adapting the operational conditions to the task at hand.

The inherent accuracy and flexibility of airborne remote sensing is currently not used because the standard method of georeferencing airborne images by available ground control limits not only the accuracy, but also often puts operational constraints on a specific flight mission. The objective of the research currently conducted at The University of Calgary is to remove these constraints and to replace the indirect method of georeferencing by ground control by a direct method of georeferencing from the aircraft. In other words, the exterior orientation of each image will be determined in real-time by onboard sensors and not in post mission by interpolation between available ground control. The fact that the method is in principle independent of available ground control has obvious economic advantages, especially in areas with poor or sparse control. The fact that position and attitude are available in real time is of no immediate advantage in current applications but may be of importance in the future.

In this paper, requirements for an airborne system of this type are presented in four points. First, the accuracies required in different application areas are discussed. Second, the performance of currently available remote sensing devices are reviewed. Third, the georeferencing problem is briefly presented, i.e. the transformation of the airborne measurements to ground level. Fourth, the implementation of the georeferencing problem by currently available position and attitude sensors is analysed. Finally, some possible sensor configurations are studied and first results of a prototype system, developed at the U of C, are presented.

The current system covers a large range of possible applications and will have specific advantages in applications of digital remote sensors, such as pushbroom scanners and CCD frame images.

2. ACCURACY REQUIREMENTS FOR DIFFERENT APPLICATION AREAS

Airborne remote sensing in its classical form of airborne photogrammetry has been widely used for cartographic mapping at all scales and currently is the only system used for high accuracy applications. Its major drawback is that the data collection process is film-based and not digital. More recently, other airborne remote sensing devices, such as pushbroom scanners, have been extensively used in those agricultural and forestry applications where accuracy requirements are not that stringent. With the ongoing improvement of scanning systems and linear array systems, a fully digital system with onboard exterior orientation, suitable for a wide range of applications, seems to be within reach. To define the design parameters of such a system, user requirements will first be discussed.

In high precision photogrammetric applications such as highway planning, large engineering projects, and cadastral applications, positional accuracies of 10 cm or less are required. In such applications, non-standard photographic overlaps of 80 % longitudinal and 60 % lateral are often employed to increase the image per object point ratio. Typical accuracy requirements at photo scales of 1:3 000 to 1:6 000 are 5-10 cm in position and 15-30 arcseconds in attitude. External attitude is not needed when a photogrammetric block adjustment approach is applied. In those cases, the geometric strength of interlocking bundles can be used to eliminate the attitude requirement.

Map accuracy is a function of map scale, thickness of printed lines, and positional accuracy of lines and points on the map sheet. The thinnest line that is legible on maps is about 0.05 mm. This high accuracy is impossible to achieve in the map compilation process which adds to the error budget through generalization, project transformation, information transfer, fine drawing, printing, etc. Typical values for the accuracy of a printed map that includes these compilation errors are five to ten times higher, see for instance Merchant (1987) and Tobler (1988). Using a line accuracy of 0.25 mm and considering that typical photo scales for base maps are between 1:12 000 and 1:15 000, the positional accuracy required of objects on the ground is 2-5m to meet the high accuracy end of cartographic applications from 1:10 000 upward.

In the resource sector, accuracy requirements cover a rather broad spectrum. At the high accuracy end of these applications, requirements are almost as stringent as in the engineering and cadastral applications mentioned above. Although sample plots with a size of 20 m by 20 m are standard in conventional forest inventory, substantially higher spatial resolution is required to reflect the internal variability of the sample stand. Detailed measurements for the tree diameter at breast height, canopy density, height etc. make a spatial resolution down to 0.25 m desirable, see Till et al (1987) for details. To discern and interpret individual trees implies a spatial resolution of less than 1 m. For damage assessment, a resolution of 5 m is needed.

Table 1 shows a summary of the accuracies required for different application areas, expressed as root mean square errors (rms) for position and attitude. It indicates that, except for a small number of high accuracy applications which require positions at the decimeter level, an accuracy of 2-5 m is fully sufficient for the bulk of the applications. This result is important for the design of GIS data bases. Instead of mixing information from different application areas with different spatial resolution requirements, it seems advisable to require a uniform spatial resolution of 2-5 m for the standard resource data base. High accuracy applications which are usually restricted to smaller projects will normally not be part of these data bases.

Application Area	RMS Accuracy for	
	Position	Attitude
Engineering, Cadastral	0.05 - 0.1 m	(15" - 30")
Cartographic Mapping 1:10 000	2-5 m	10' - 20'
Resource Applications	2 - 5 m	20' - 30'
Forestry (Detailed)	0.2 - 1.0 m	1' - 3'

Table 1: Accuracy Requirements

3. ACCURACY OF CURRENT REMOTE SENSORS

The georeferencing requirements of an airborne positioning and attitude system is determined by the spatial resolution of the remote sensor. Commonly used sensors such as photographic systems, scanning and linear array systems, and synthetic aperture radar (SAR) have quantifiable spatial resolution limitations. As such the following discussion will be focused on these sensor types.

3.1 Spatial Resolution of Photographic Systems

The spatial resolution (R) of an aerial photograph is influenced by a number of factors such as the resolving power of the camera lens and the film used in a photographic system. In addition, the spatial resolution is affected by any uncompensated image motion during exposure, the atmospheric conditions present at the time of image exposure, and the conditions of image processing (Lillesand and Kiefer, 1987). Also, the focal length (f) and the distance (d) between a target and the camera also determine the spatial resolution of a photograph. Among these, only the resolving power of the photographic system and the uncompensated image motion may be quantifiable. The resolving power of a photographic system is expressed in number of line-pairs/mm (n) (i.e., black and white line pairs of equal thickness (Wolf, 1974). The optical quality of the lens, the granularity and the speed of the film all contribute to the determination of the resolving power of a photographic system. Under a range of contrast of black and white between 2:1 to 1000:1, the resolving power of photographic systems ranges from 50 line pairs/mm to 100 line pairs/mm (Lillesand and Kiefer, 1987). Due to a number of other factors mentioned above, R is usually poorer than d/(fn2000) m. Thus,

$$R > d/(fn2000) \text{ m}. \quad (1)$$

The ratio f/d determines the local image scale (s) for the photographed target. For example, an aerial photograph with a 1:10000 image scale and a resolving power of 50 linepairs/mm, has a spatial resolution (R) which is 0.1 m or less. Consequently, the positional accuracy requirements are 10 cm while the attitude requirements correspond to 15 arcseconds for the exterior orientation of an individual photograph.

In general, multiple strips of photographs are used in a block adjustment to obtain a favorable error distribution making use of the inherent geometrical strength of the photographic image. In this case, georeferencing can be done by position control only. Precise independent attitude is not needed because bundles of interlocking rays will take care of this requirement. By accurately fixing the perspective centres of these bundles in space, even high accuracy requirements can be met.

3.2 Spatial Resolution of Scanning Systems and CCD Frame Imagers

Compared with a photographic system, the only influencing factor that is different in a scanning system or a CCD frame imager is that the resolving power of the film has been replaced by the size (z) of charge-coupled devices (CCDs). Since the resolving power of a camera lens is considerably higher than the size of a CCD, the determining factor becomes the size of the CCD. Similar to the photographic systems, the spatial resolution (R) for a sensor system based on CCD technology cannot be better than z/s or dz/f, i.e.

$$R > dz/f, \quad (2a)$$

when f, d, and z are given. For CCD-based airborne sensors, often the physical dimension (p) of the CCD array, the number of CCD elements in a line (nc) and the camera field-of-view angle (ß) are specified. Here, f can approximately be obtained ptan(ß/2)/2. Similarly, the resolving power z is given by p/nc. Therefore, with substitution,

$$R > (2d \tan ß/2)/nc \quad (2b)$$

For example, a CCD camera with a 28 mm lens and a CCD array of dimension 30 mm by 30 mm with 4K by 4K CCD elements. When used at a flying height of 1000 m, R> 0.25 m. Increasing the focal length or the number of CCD elements per mm, will result in increased resolution and increased storage rates. For the example given above, the storage rate is 1.6 MB per second at an aircraft speed of 100 knots and 50% overlap. This is the limit for current hard disk technology and a major reason why high resolution CCDs have not been used in airborne applications.

If the aircraft flight speed is 100 knots, and the shutter speed is 1/250 seconds, the smearing caused by aircraft movement without any compensation is 0.1 m at the ground level. This will degrade the spatial resolution of sensor systems along the flight direction, but can be minimized by using high accuracy velocity output of the georeferencing sensors. Because CCD cameras do not have the geometrical properties of aerial photographs, both attitude and position are needed for georeferencing pushbroom scanners and linear array systems. Whether or not motion compensation is necessary, depends on the accuracy requirements of a specific application.

3.3 SAR Accuracy Requirements

Synthetic-aperture radar (SAR) is different from the sensors discussed so far because it requires high precision velocity input in real time to compensate for aircraft motion. In particular, the SAR imagery is realized as a result of an accumulation of several 'looks' which are acquired while the aircraft is in motion. The accumulated signal is used to minimize the inherent speckle of radar imagery. Since the velocity requirements are very stringent, it is often overlooked that the final accuracy of the radar map depends heavily on the georeferencing accuracy. Currently, georeferencing is done by interpolating between GPS ground control and by matching visually identified features.

Design specifications will be taken from the STAR-1 system because it is currently the only system that operates commercially. The operational parameters of the STAR-1 system include a nominal flying altitude of about 10 km to provide swath coverage and high fuel efficiency. Operating in the X-band of the electromagnetic spectrum, the STAR-1 system can map on either side of the aircraft with a 90° orientation to the flight path. The corresponding azimuth resolutions of the wide swath (WS) and high resolution (HR) modes are 12 m and 6 m, respectively. The radar may use up to seven looks in azimuth by one in range per swath in both modes. It is designed for a maximum ground speed of 350 knots. The pulse width is 30 microseconds and is operated at a pulse repetition frequency (PRF) of 1200 Hz with a maximum duty cycle of 0.036 (repeat time). Correspondingly, velocity measurements with an accuracy of 0.0002-0.0005 m/sec are required to precisely correct the multi-look averaging process.

While positional accuracies of current radar maps are at the level of 15-20 m, georeferencing requirements should be based upon a 6 m resolution and a seven look sampling. This implies that accuracy requirements in position are about 2-4 m. If independent attitude is needed for each swath, the corresponding accuracy would be 10-40 arcseconds. Usually, attitude requirements are not that stringent because differential attitude is derived from the radar ranges.

Recently, methods have been proposed to use SAR interferometry for topographic height determination. In this case, pairs of SAR images are used which have been acquired from the same sensor flown on parallel flight lines at approximately the same flying height. Instead of using a radargrammetric approach (Leberl, 1990) for height determination, the phase differences of the SAR image pairs are used to produce interferograms which form the basis for topographic height determination. Such radar interferometric techniques use the range difference information in the form of phase differences in conjunction with the flight path parameters to resolve the terrain height (Madsen et al, 1993). Given similar single image resolution specifications as with SAR systems mentioned above, the required positional accuracies would be about 2-4 m. Due to the use of dual ranges, the attitude accuracy requirements would be reduced if relative heights were required. The determination of absolute terrain heights to 2 m would require attitude accuracies of about 10-40 arcseconds. SAR systems need position for georeferencing, and precise velocity for motion compensation in real time. Attitude is required in radar interferometry but may be optional and radargrammetry.

3.4 Summary of Results

Table 2 summarizes the highest georeferencing requirements for the three types of remote sensing systems discussed above. It indicates that they are ordered in terms of increasing complexity with respect to the georeferencing requirements. While accurate positioning is usually sufficient for photogrammetric sensors, position and attitude is required for pushbroom scanners and CCD frame imagers, and position, attitude, and velocity is needed for SAR systems.

Type of Sensor	Georeferencing Accuracy Required: Position	Attitude	Velocity
Photo Camera (low flying altitude)	0.05-0.1 m	(15"-30")	not requ.
Digital Scanning Systems (CCD)	0.25-1.0 m	1'-3'	1-2 cm/s
SAR Systems	2-4 m	(10"-40")	0.02-0.05 cm/s

Table 2: Imaging Sensor Accuracies

4. GEOREFERENCING MODEL

Georeferencing describes a series of transformations necessary to obtain coordinates in a chosen mapping system (m) from the output of a remote sensing device in the body frame (b) of the aircraft. The major steps in this transformation are depicted in Figure 1, Schwarz et al (1993).

The mathematical model corresponding to this figure is

$$\Delta r^m(t) = r^m(t) + s_m R_b^m(t)\, p^b \quad , \qquad (3)$$

where

Δr^m is the position vector of an image object in the chosen mapping frame;

r^m is the coordinate vector from the origin of the mapping frame to the centre of the position sensor on the airplane, given in the m-frame;

R_b^m is the three-dimensional transformation matrix which rotates the aircraft body frame into the mapping frame (roll, pitch, and yaw are measured by the INS);

s is a scale factor derived from the height of the sensor above ground;

p^b is the vector of image coordinates given in the b-frame.

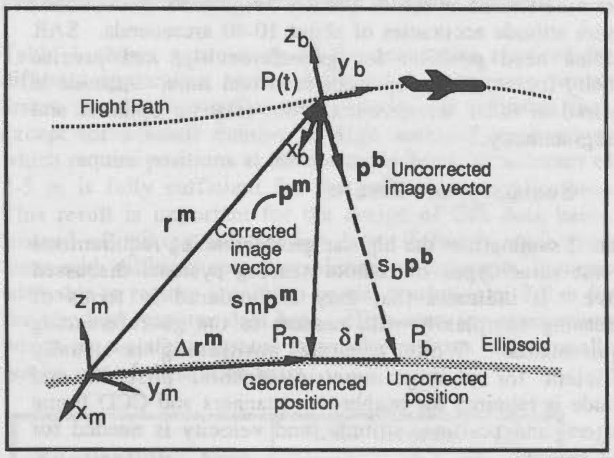


Figure 1: Georeferencing of Airborne Sensing Data

Equation (3) is, however, only a first approximation of the actual situation. The three sensors for positioning, attitude determination, and imaging are physically separated, and it can therefore not be assumed that they function in the same measurement frame. The actual situation is shown in Figure 2 which enlarges the area around point P(t) in Figure 1. It has been assumed that the remote sensor, for example a photogrammetric camera, is mounted in the stable area of the airplane, that the positioning sensor, a GPS antenna is mounted on top of the airplane, and that the attitude sensor, an inertial measuring unit is mounted in the interior of the aircraft, somewhere close to the remote sensor. In this case, aircraft position is defined by the antenna centre of the GPS receiver (m-frame) and aircraft attitude is given by the internal axes of the inertial measuring unit (b-frame).

They do in general not correspond to the position and attitude of the remote sensing device which is given by the position and orientation of the camera frame (c-frame). This frame has its origin in the perspective centre of the camera, its z-axis is defined by the vector of length f between the perspective centre and the principal point of the photograph, and its (x,y)-axes are defined in the plane of the photograph and are measured with respect to the principal point. The corresponding image vector is therefore of the form

$$p^c = \begin{pmatrix} x - x_p \\ y - y_p \\ -f \end{pmatrix} \qquad (4)$$

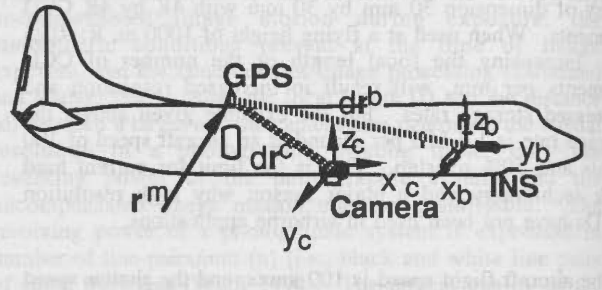


Figure 2: Coordinate Transformations Between Airborne Sensors

In case of pushbroom scanners and CCD fram imagers, the second vector component is replaced by

$$y^c = (y - y_p) / ky$$

where ky accounts for the non-squareness of the CCD pixels. The resulting modelling equations are

$$\Delta r^m(t) = r^m(t) + R_b^m(t) \{ s\, dR_c^b\, p^c - dr^b \} \qquad (5)$$

where the subscripts and superscripts correspond to the frames defined above. The additional notations in Equation (1) are as follows:

dR_c^b is the transformation matrix which rotates the camera frame into the body frame;

p^c is the imaging vector in the c-frame as given by Equation (4)

dr^b is the translation vector between the GPS antenna centre and the centre of the INS, and

dr^c is the translation vector between the GPS antenna centre and the perspective centre of the camera.

This equation, in a somewhat simplified form, has been discussed in detail in Schwarz et al (1993). A few remarks will therefore suffice here. It should be noted that the origins of the position and attitude sensors are not identical. Furthermore, the vectors r^m and Δr^m, as well as the rotation matrix R_b^m are time dependent quantities while the vectors p^c and dr^b as well as the matrix dR_c^b are not. This implies that the aircraft is considered as a rigid body whose rotational and translational dynamics is adequately described by changes in Δr^m and R_b^m. This means that the translational and rotational dynamics at the three sensor locations is uniform, in other words, differential rotations and translations between the three locations as functions of time have not been modelled. It also means that the origin and orientation of the three sensor systems can be considered fixed for the duration of the flight. These are valid assumptions in most cases but may not always be true.

The quantities Δr^m , R_b^m and p^c in Equation (1) are determined by measurement, the first two in real time, the third in post mission. The quantities dR_c^b and dr^b, however, are determined by calibration, either before or during the mission; for details see Schwarz et al (1993). To define dR_c^b by calibration, a minimum of three well determined ground control points is required. The scale factor s is changing with the flying altitude of the aircraft above ground. It can, therefore, either be approximated by assuming a constant flying altitude, calibrated by introducing a digital terrain model, or determined by measurement, using either stereo techniques or an auxiliary device such as a laser scanner. For precise georeferencing, the latter techniques are the most interesting to be investigated because they would provide all necessary measurements from the same airborne platform and thus avoid datum problems.

The above equation can be used to evaluate the georeferencing requirements for photographic systems, scanning systems, CCD fram imagers, and radargrammetric systems. The overall accuracy will depend on the resolution of the remote sensing device and the accuracy with which the parameters in Equation (5) can be determined. The important parameters are the accuracy of the position and attitude determination on the one hand and the stability of the sensor configuration on the other. This will be further investigated in the next section. It should be noted, however, that for radargrammetric systems, velocity is an additional parameter which has to be determined with high accuracy. It is required for motion compensation which strictly speaking is not part of the georeferencing process but part of the remote sensing process and therefore has to be accomplished in real-time. Since the position and attitude sensors discussed in section 5 will provide velocity as a by-product, it will be included in the following discussion.

5. PERFORMANCE OF POSITION AND ATTITUDE SENSORS

To achieve the required accuracies in position and attitude, two major systems are currently available, GPS and INS. GPS is primarily a positioning device, measuring distances to satellites whose positions are known. It can be used as an attitude sensor, however, by transforming vector changes in a fixed antenna configuration into attitude changes. INS has two independent sensor triples to measure accelerations and angular velocities from which linear velocity, position, and attitude can then be derived by integration.

The two systems have very different error characteristics which are due to the type of measurements used. GPS accuracies are essentially uniform and time independent. Variations in accuracy are mainly due to satellite configuration and atmospheric conditions. The error spectrum for position is essentially flat and more or less stationary. INS accuracies are heavily affected by the fact that all measurements have to be integrated to obtain the required position and attitude parameters. Since the error spectrum is not flat but shows some low frequency spectral lines, position and attitude accuracies deteriorate in a systematic manner as a function of time. Thus, short term accuracy is excellent and equivalent or better than GPS accuracy, long term accuracy is not and needs updating to stay in the range required for precise georeferencing.
From an operational point of view, the higher output rate of inertial systems (typically 50-100 Hz) is a major advantage because the exterior orientation of each image or each scan line can be determined without interpolation or prediction. Current GPS output rates (typically at 2 Hz, with emerging systems at 10 Hz) will not allow direct computation.

5.1 GPS Performance

Table 3 summarizes the positioning and attitude accuracies that are currently achievable using GPS. The single point positioning error budget is dominated by Selective Availability (SA), especially satellite clock dithering. The achievable RMS accuracy quoted in the table, i.e. 100 m horizontal and 150 m vertical, is therefore essentially independent of the type of receiver used. Although they are not shown in the table, velocity errors are also affected by SA. An RMS accuracy of about 0.5 m/s can be achieved in single point mode.

Model	Accuracy	
Pseudo range point positioning*	100 m horizontal 150 m vertical	
Smoothed pseudorange differential positioning	10 km	0.5 - 3 m horizontal 0.8 - 4 m vertical
	500 km	3 - 7 m horizontal 4 - 8 m vertical
Carrier phase differential positioning	10 km	3 - 20 cm horizontal 5 - 30 cm vertical
	50 km	15 - 30 cm horizontal 20 - 40 cm vertical
	200 km (with precise orbits, same as 50 km)	
Attitude determination	1 m separation	10-30 arcminutes
	5 m separation	4 - 6 arcminutes
	10 m separation	2-3 arcminutes

*Selective Availability on, PDOP ≤ 3, 2DRMS (95%) (DOT/DOD, 1992)

Table 3: GPS Positioning and Attitude Accuracies

For DGPS positioning at the decimeter level, the carrier phase observable is required. Once the carrier phase integer ambiguities are resolved, the positioning accuracy is at the level of a few centimetres to decimeters for separations of less than 10 km, depending on multipath and atmospheric effects, see Table 4. The tropospheric effect may be rather significant for airborne positioning as reported in Tiemeyer et al. (1994). As in the smoothed pseudorange case, the orbital and atmospheric errors decorrelate with an increase in the monitor-remote separation at the level of 2-5 ppm. Shi and Cannon (1994) presents airborne results at the few decimeter level for baselines up to 200 km when dual frequency receivers were used in conjunction with precise orbits.

When using the differential mode of operation (DGPS), receiver characteristics become very important in the determination of the level of performance. Table 1 shows that RMS accuracies of 0.3 - 3 m can be reached for the horizontal component while 0.5-4 m can be achieved vertically using the smoothed pseudo range DGPS model. The range takes into account the noise of the measured pseudo range as well the receiver's susceptibility to multipath, which are the two dominant errors for a 10 km monitor-remote separation. Several tests have confirmed values at the high accuracy end of the range when narrow-correlator C/A code receivers or P code receivers are used (e.g. Cannon et al., 1992). Standard C/A code receiver technology typically gives DGPS accuracies at the upper end of the range. The accuracy degrades as the separation between the monitor and remote receivers increase which is due to additional errors from the orbit and atmosphere. All results given are for post-mission operation. Real-time operation is possible with an appropriate data link but accuracies are typically degraded with respect to those given in Table 3.

The use of GPS for attitude determination can take on a variety of forms ranging from a self-contained system which has a number of channels divided between several antennas, to an independent system which is made up of individual antenna/receivers. The second option is flexible in the sense that the receivers can be used for a variety of applications in addition to attitude determination. A minimum of two antennas is required for heading determination while at least three are needed to obtain roll, pitch and azimuth. Redundant antennas are also generally used to improve system reliability. The achievable accuracy is mainly a function of baseline length between antenna pairs, i.e. the longer the separation the higher the accuracy. Table 1 gives the level of accuracy as a function of separation which is usually determined by platform limitations. Many land-based applications require shorter baselines, while airborne and marine platforms can tolerate separations of the order of 5 - 10 m. A number of tests have been conducted with multi-antenna systems and accuracies which fall within the levels presented in Table 1 have been demonstrated, see e.g. Schwarz and El-Mowafy (1992) for a comparison with INS and Schade et al. (1993) for a comparison derived from inverse photogrammetry. The accuracies given in the table may be somewhat conservative. Recently, El-Mowafy and Schwarz (1994) reported accuracies of 3 arc minutes (RMS) for a 3m baseline.

5.2 INS Accuracy

Inertial sensors are devices sensing either linear acceleration or angular velocity. When assembled into an inertial navigation unit, they instrument an autonomous system for three-dimensional velocity, position, and attitude determination. In the following, only such systems, not individual inertial sensors, will be discussed.

Inertial navigation systems come in two major varieties, namely as stable platform systems and as strapdown systems. Stable platform systems establish the internal attitude reference mechanically and provide a platform orientation with respect to some prescribed coordinate system. In strapdown systems, the same process is done analytically, i.e. angular velocities with respect to the body frame are measured at a high rate and orientation changes with respect to the prescribed frame are computed. Although both types of systems can be used for the problem at hand, strapdown systems have major advantages in terms of data rate, attitude output, failure rate, price, power requirements, and weight. The following discussion will therefore be directed towards systems of this type.

In georeferencing applications, inertial systems function mainly as precise attitude systems and as short-term interpolators for velocity, position, and attitude. Because of the time dependence of all major errors, the long-term velocity and position performance is not sufficient for precise georeferencing. Thus, regular position and/or velocity updates are needed to keep the overall errors within prescribed boundaries. The accuracy of inertial systems depends heavily on the quality of the sensors used which themselves are a function of the system costs. We will therefore distinguish between high accuracy, medium accuracy, and low accuracy systems.

Their error characteristics are summarized in Table 4 using four typical time intervals. The one hour interval characterizes the long-term behaviour and is therefore an indicator of the suitability of using an INS as a stand-alone georeferencing system. The one minute interval characterizes the short-term interpolation accuracy, including bridging for GPS outages, and GPS cycle slip detection and fixing. The one second interval characterizes the interpolation time for an integrated GPS/INS, assuming that gyro drift can be eliminated. The attitude accuracies given are for pitch and roll, those for heading are three to five times larger. Noise levels in attitude can vary considerably, depending on the type of gyros used. In general, dithered ring-laser gyros have higher noise levels if the compensation loops have not been specifically designed for high short-term attitude and velocity.

Table 4 shows that only high accuracy inertial systems can be used for stand-alone georeferencing, and that even in that case, the positioning accuracy is marginal. On the other hand, in the short term, the attitude accuracy is far superior to that obtained from GPS multi-antenna systems. Thus, for short-term interpolation of position and attitude, both the high accuracy and the medium accuracy inertial system are suitable. Both will also provide the velocity accuracies needed for motion compensation in SAR systems. It appears therefore that an integrated GPS/INS provides the best guarantee for a georeferencing system that will cover a wide range of applications.

Error in	System Accuracy Class (rms) high	medium	low
Attitude			
1 h	10" - 30"	1' - 3'	1° - 3°
1 min	1" - 2"	15" - 20"	0°.2 - 0°.3
1 s	< 1"	1" - 2"	0°.01 - 0°.03
50Hz (noise)	0".1 - 0".2	0".1 - 0".2	15" - 20"
Velocity			
1h	0.3 - 0.5 m/s	1 - 2 m/s	200 - 300 m/s
1 min	0.01 - 0.02 m/s	0.05 - 0.1 m/s	1 - 2 m/s
1 s	0.0005 - 0.001 m/s	0.001 - 0.003 m/s	0.002 - 0.005 m/s
50 Hz (noise)	0.0002 - 0.0005 m/s	0.0005 - 0.002 m/s	0.001 - 0.003 m/s
Position			
1 h	0.3 - 0.5 km	1 - 3 km	200 - 300 km
1 min	0.3 - 0.5 m	0.5-3.0 m	30 - 50 m
1 s	0.01 - 0.02 m	0.03-0.10 m/	0.3 - 0.5 m
50 Hz (noise)	0.005 - 0.001	0.001-0.005 m	0.05 - 0.10 m

Table 4: INS Position, Velocity, and Attitude Accuracies

Accuracy Required	System Configuration	Airborne Sensor	Cost (K$=$1000)	Characteristics
0.05 - 0.1 m 15" - 30" 0.0002 - 0.0005 m/s	High accuracy INS plus DGPS (carrier phase)	All airborne sensors	K$250	Suitable for all applications
0.05 - 0.1 m	DGPS (carrier phase)	Photogrammetric camera	K$50	Block adjustment only
2 - 5 m ≥ 10' 0.01 - 0.02 m/s	DGPS (pseudo-range) plus GPS multi-antenna system	Pushbroom scanner, CCD frame imagers	K$50	Low data rate. Attitude transfer to imaging sensor is problematic (stability)
2 - 5 m ≥ 1' 0.0002 - 0.0005 m/s	DGPS (pseudo-range) plus medium accuracy INS	Pushbroom scanners, CCD frame imagers, some SAR applications	K$100 - 120	Long-term velocity may be marginal

Table 5: Possible Sensor Configurations

6. POSSIBLE SENSOR CONFIGURATIONS

In Table 5, the possible sensor configurations for each accuracy range are given and some advantages and drawbacks are listed. A cost estimate based on current hardware costs is also attached. It does not include interfacing of the sensors and dedicated software development which will be necessary in most cases. It also does not include aircraft installation and certification which usually adds considerably to the costs. Considering, however, that such a system would eliminate all need for ground control, except the minimum required for calibration, the overall costs are reasonable.

Table 5 shows that it is possible to develop a georeferencing system for airborne remote sensing that meets all current accuracy requirements by integrating available INS and GPS hardware components. To achieve the position accuracy for high precision engineering and cadastral applications, the distance to the remote receiver should not be more than 15-20 km. The only potential error source which has not been extensively studied are differential rotations between the INS and the remote sensing device. In other words, to achieve the high attitude accuracy, flexing of the aircraft hull, as well as vibration differences at the two locations have to be investigated. In principle, this amounts to the question whether or not the rotation matrix dR_c^b can be considered as time invariant or not. If the differential rotations are well below the noise level of 15"-30", current high precision INS together with DGPS can be used to meet the attitude requirements of even the most demanding applications discussed in this paper.

The determination of the translation vector dr^b is not as critical. Even if the length of the vector changes somewhat in time, it will usually stay well within the accuracy level of 5-10 cm. If photogrammetric cameras are used for high precision applications and a block adjustment is possible, the georeferencing problem can be reduced to precise positioning of the aircraft. This can be done with sufficient accuracy by using differential GPS and carrier phase techniques. In this case, the costs for the georeferencing system can be considerably reduced.

If high accuracy applications are excluded, more economic solutions to the georeferencing problem can be found. Such systems would be suitable for the bulk of applications in the resource sector. They will therefore most likely be the standard georeferencing systems of the future.Two solutions seem to be especially attractive. One is completely based on GPS technology, the other is a GPS/INS integration with a lower cost INS.

The GPS solution would combine two narrow-correlator C/A code receivers for positioning with a GPS multi-antenna system for attitude determination. The system would be low cost and has the advantage that the same receivers can be used for both tasks and sufficient redundancy can be built into the system design. Drawbacks are system stability and data rate. To get full attitude resolution, one or two of the antennas have to be installed on the wings. The resulting attitude is affected by wing flutter and most likely deviates considerably from the attitude of the remote sensing system. This problem is aggravated by the fact that a 50 Hz output rate is needed. Current GPS output rates are at 2 Hz, thus interpolation is needed. The compounded influence of the two effects has to be investigated to confirm the suitability of such a system for the stated accuracy requirements. The system would not be suitable for SAR-type applications because the short-term velocity accuracy is not sufficient. The medium cost GPS/INS integration solves the above problems and has sufficient short-term velocity resolution to also be used in SAR applications. It covers therefore a rather broad range of applications. Since the INS can be mounted on the same frame or the same platform as the imaging system, there is no problem with system stability, as long as the rigidity of the frame under vibrations is carefully checked. Output rate is not a problem because inertial systems come with rates of 50-100 Hz anyway. A further advantage is that an onboard INS can usually be modified to form part of the integrated system. This will reduce the costs considerably. It appears therefore that the development of such a system offers the best balance between economy and technological risk. This is one of the major reasons why the prototype development at the U of C moved in this direction.

7. TESTING OF THE U OF C PROTOTYPE SYSTEM

The current prototype system at the U of C consists of a strapdown INS of the medium accuracy class (Litton LTN 90/100) and two Ashtech P XII receivers. This system was recently tested with the Compact Airborne Spectrographic Imager (***casi***) developed by Itres Research Ltd for this project (Babey & Anger, 1989). The casi is a pushbroom sensor that acquires one scanline at a time as it travels along the flight line. The resulting image possesses a different set of position and orientation parameters for each scanline, and it often contains large distortions induced by movements on the aircraft. An example of such an image is shown in Figure 3.

The calibration parameters in Equation (5) are f, x_p, y_p, ky, and dR_c^b and dr^b. All parameters, except dr^b were determined from a calibration target array in the area flown, using a self-calibrating bundle adjustment. The orientation differences between INS and the imaging sensor, dR_c^b will change each time the imager and INS is installed in the aircraft. In fact, dR_c^b may change during flight if the camera is not rigidly secured with respect to the INS. However, dR_c^b can be calibrated from the imagery if three or more control points are present in the block. A block is defined by a series of overlapping flightlines. For convenience four points were chosen for calibration. A bundle adjustment is used to solve for these parameters, Gibson & Buchheit (1990).

To fully geocorrect the imagery, Cosandier et al, (1992) and create an ortho-image, the ground height must be taken into account via a digital elevation model (DEM). The integration and extraction of a DEM from ***casi*** is currently being developed.

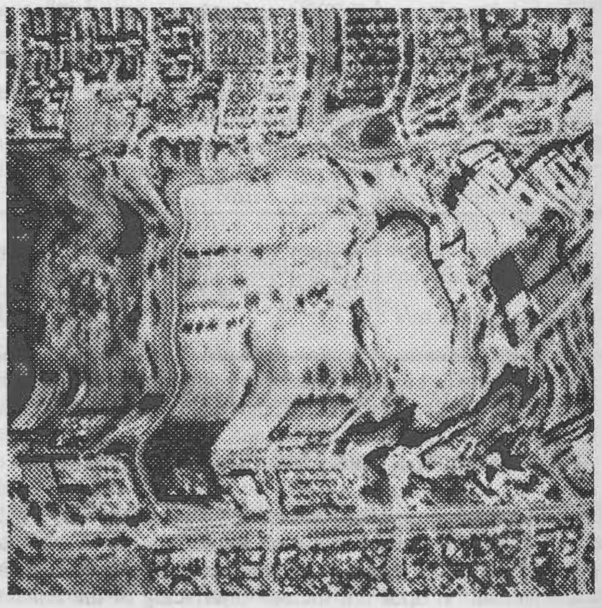


Figure 3: Raw Pushbroom Image

7.1 Calgary Test Flights

A special variant of the *casi* was developed by Itres Research Ltd. for this project. This sensor contains fore, nadir and aft look directions giving it stereo capabilities. The sensor has 512 pixels along each scanline and a field of view of 43.5°. The fore and aft look directions are 16° off nadir. These limited off-nadir angles were due to the CCD geometry.

Three flightlines were flown over the city of Calgary. Each flight line was 10,000 lines long with an along track pixel size of 1.6 metres and an across track pixel size of 2.5 metres. Included on board the aircraft were the *casi* sensor, a Litton LTN-90 inertial system and an Ashtech P12 receiver. Another GPS receiver was placed near the centre of the test area. GPS positions were processed to 0.50 metre accuracy.

Due to the fairly large pixel size, large area and urban nature of the area, placing targets on the ground was impractical. Therefore, road intersection centres were used as targets instead. These control points are somewhat poor due to errors in matching the road intersection centre on the ground to the image. However, in a production environment, this would be the most likely way control would be established.

7.2 Test Results

To test the accuracy of the system, bundle adjustment results are shown. These would be equivalent to geocorrection accuracies if the DEM was taken into account. Each adjustment solves for a different dR_c^b for each flight line. The values for the orientation differences ω, ϕ and κ are compared for each flight line.

Three different solutions are shown for the above mentioned data set. They are:

a) First, the bundle adjustment was run with all control points and fore, aft and nadir look directions. This gives the optimal solution from all available data and is used to compare results.

b) Next, the bundle adjustment was run with only 4 control points and fore, aft and nadir look directions. This corresponds to the case of minimum ground control for calibration. The remaining control points are included as check points. This would be a standard ***stereo*** scenario.

c) Lastly, the bundle adjustment was run with 4 control points and only the nadir look direction. The remaining control points are included as check points. This scenario is would be used with the standard multi-spectral ***casi***.

Test (a) results show that the overall horizontal accuracy is about 1.2 pixels (3.0 metres). The largest control point error was 5 metres *(horizontal)*. Larger errors in the Z axis are due partially to a poorer height geometry as seen by the standard errors, and also due to poor control point registration between the image and ground.

	X	Y	Z	no.
Control points. RMS (m)	2.778	2.861	14.88	27
Check points RMS (m)				
Image residuals (pixels)	1.00	0.80		1057
Avg. standard error (m)	2.32	2.11	8.60	281

Table 6: Test *(a)* Results, All Control Points Used With Stereo Pushbroom Sensor

With improved targeting, the control point residuals will be reduced as they are currently the largest error source. Other error sources result from the GPS and INS, but these are both below the sub-pixel level. There are also errors due to the calibration, but they are difficult to quantify individually. The principal point (≈CCD centre) is most difficult to calibrate. However, erroneously calibrated parameters will be partially compensated by the INS/sensor offset angles.

	ω (deg.)	ϕ (deg.)	κ (deg.)
Flightline 1	1.9697	5.7278	3.1517
Flightline 2	1.9203	5.5642	3.4284
Flightline 3	1.9998	5.7948	2.7574

Table 7: INS/Sensor Offset Angles for Test *(a)*

The values for the INS/imaging sensor orientation offset dR_c^b are shown in Table 7. Although they are relatively consistent, there are some variations present. Most significant variations are seen in the κ (yaw) values. This larger variation is expected because the system, with its fairly narrow swath angle (43.5°), is less sensitive to yaw. Other differences may also be due to the fact that flightlines 1 and 3 were in one direction and flight line 2 in the opposite direction.

	X	Y	Z	no.
Control points. RMS (m)	1.497	1.401	2.325	4
Check points RMS (m)	2.646	3.701	22.33	23
Image residuals (pixels)	0.99	0.72		1057
Avg. standard error (m)	2.43	2.18	8.67	301

Table 8: Test *(b*)Results, 4 Control Points Used With Stereo Pushbroom Sensor

Table 8 shows results of test (b) using only 4 control points. These results are very similar to those of Table 6 which means there is no significant accuracy degradation due to using fewer control points. Four points are sufficient for this size of block. This means that after calibration, the georeferencing system performs with the expected accuracy.

Tables 9 and 10 show results using only the nadir channel. This is the geometry of the standard ***casi*** sensor. The errors in y are slightly larger than from the *(a*)and *(b*)tests.

	X	Y	Z	no.
Control points. RMS (m)	0.699	0.310	0.109	4
Check points RMS (m)	2.748	4.791		9
Image residuals (pixels)	0.11	0.61		348
Avg. standard error (m)	2.43	2.18	8.67	47

Table 9: Test (*c*)Results, 4 Control Points Used With Nadir Only Pushbroom Sensor

	ω (deg.)	φ (deg.)	κ (deg.)
Flightline 1	2.1025	5.7231	3.3090
Flightline 2	1.7699	5.4316	3.1400
Flightline 3	2.0822	5.7064	2.6167

Table 10: INS/Sensor Offset Angles for Test (*c*)

This is due to a poorer geometry with the absence of fore and aft channels. However, the image residuals are very low indicating a good internal consistency. Finally, the distorted image contained in Figure 3 is geocorrected and shown in Figure 4.

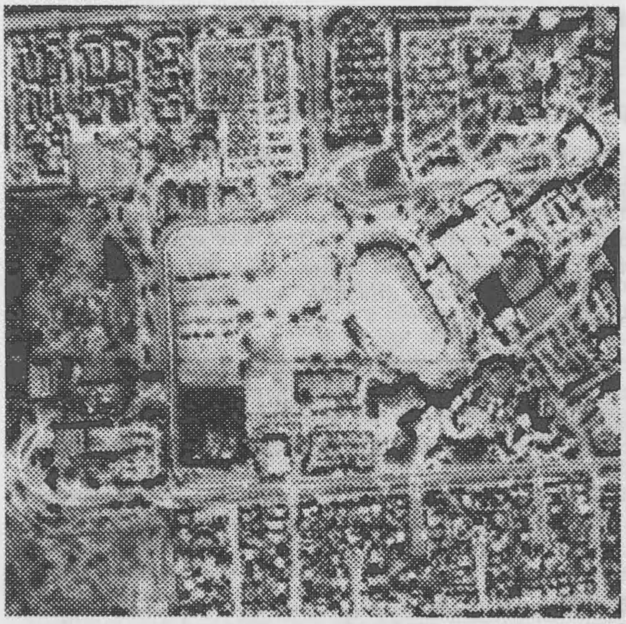


Figure 4: Geocorrected image

Overall, the test results show that georeferencing by airborne sensors is possible for pushbroom scanners and CCD frame imagers of this accuracy class. The size of the residuals at the check points seems mainly due to the quality of the ground control (2-3 m), and the pixel size. Georeferencing errors seem to play a minor role at this level of accuracy. Additional tests using airborne photogrammetry and a precisely controlled field of ground targets are planned in the near future. They will provide a reliable estimate of current system accuracies.

CONCLUSIONS

Results presented in this paper indicate that georeferencing of airborne remote sensing data is possible by an onboard GPS/INS system using currently available off-the-shelf hardware. In this system, GPS provides accurate position information, while INS provides precise attitude and velocity information. If a high precision INS is used as a component of the integrated system, all current accuracy requirements, including those for high precision engineering and cadastral applications, can be met. If a system for the somewhat lower accuracy requirements in the resource sector are needed, a medium accuracy INS can be used and the total hardware costs can be reduced to one half. A georeferencing system based on GPS technology only, needs further testing before its suitability for these applications could be affirmed.

Results of a first series of tests with the casi system indicate that the georeferencing requirements for this pushbroom scanner are met if four ground control points are used for system calibration. It appears that the residual errors on check points are dominated by errors in the control points (non-targetted) and the pixel size. A more stringent test to assess the capability of the current system using photogrammetric techniques is planned in the near future.

ACKNOWLEDGEMENTS

Work on this project was supported by an NSERC Strategic Grant to the first four authors. Itres Research Ltd is thanked for its interest in this project and its participation in system testing. Dr. Ming Wei and Mr. Gengsheng Zhang are acknowledged for contributing to the INS specifications.

REFERENCES

Babey, S.K., C.D. Anger, 1989. A compact airborne spectrographic imager (CASI), Proceedings, International Geoscience and Remote Sensing Symposium, Vancouver, Canada, pp. 1028 - 1031.

Cannon, M.E., and G. Lachapelle, 1992, Analysis of a High Performance C/A Code GPS Receiver in Kinematic Mode, Navigation, Journal of the U.S. Institute of Navigation, Vol. 39, No.3, pp 285-299.

Cosandier, D., T. Ivanco, S. Mah, 1992. The geocorrection and integration of the global positioning system with the compact airborne spectrographic imager, 15th Canadian Symposium on Remote Sensing, June.

DOT/DOD, 1992, 1992 Federal Radionavigation Plan.DOT-VNTSC-RSPA-92-2/DOD-4650.5. National Technical Information Service, Springfield, VA.

A. El-Mowafy, and K.P. Schwarz, 1994. Epoch by Epch Ambiguity Resolution for Real-Time Attitude Determination Using a GPS Multi-Antenna System. Submitted to Navigation, Journal of the US Institute of Navigation.

Gibson, J.R., M. Buchheit, 1990. Precise Geometric Processing of Stereo MEIS Imagery, Proceedings ISPRS, Commission 7.

Leberl, F.W.,1990. Radargrammetric Image Processing, Artech House, Boston.

Lillesand, T.M., and R.W. Kiefer, 1987. Remote Sensing and Image Interpretation, Sec. Ed., John Wiley and Sons: Toronto.

Madsen, S.N., H.A. Zebker and J. Martin, 1993. Topographic Mapping using Radar Interferometry: Processing Techniques, IEEE Transactions on Geoscience and Remote Sensing, Vol. 31, No.1, January, pp. 246-256.

Merchant, D.C., 1987. Spatial accuracy specification for large scale topographic maps. PE&RS, 53(7):958-61.

Nichol, A. J. Wilhelm, R.Inkster, S.Leung and T. Gaffield, 1986. A SAR Real-Time Ice Reconnaissance, IEEE Transactions on Geoscience and Remote Sensing, Vol. GE-24, No.3, May.

Schade, H., M.E. Cannon and G. Lachapelle, 1993. An Accuracy Ananylis of Airborne Kinematic Attitude Determination with the NAVSTAR/Global Positioning System, SPN Journal, Vol. 3, No. 2, pp. 90-95.

Schwarz, K.P., A. El-Mowafy, M. Wei, 1992. Testing a GPS Attitude System in Kinematic Mode. 5th Int. Techn. Meeting ION-GPS 92, Albuqerque, N.M., Sept. 16-18, 1992. pp. 801-809.

Schwarz, K.P., M.A. Chapman, M.E. Cannon, P. Gong, 1993. An Integrated INS/GPS Approach to the Georeferencing of Remotely Sensed Data. Photogrammetric Engineering and Remote Sensing, 53, 2. pp. 1667-1674.

Shi, J. and M.E. Cannon, 1994. High Accuracy Airborne GPS Positioning Using a Multi-Receiver Configuration, Proceedings of IEEE PLANS'94, Las Vegas, April 11-15 (in press).

Tiemeyer, B., M.E. Cannon , G. Lu, and G. Schänzer, 1994. Ambiguity Resolution of GPS L1 Carrier-Wave Measurements for High Precision Aircraft Navigation, Z. Flugwiss.Weltraumforsch, Springer Verlag. (in press).

Till, S.M., R.A. Neville, D.G. Leckie, and M.W. Strome, 1987. Advanced airborne electro-optical imager. Proceedings of the 21st Symp. on Remote Sensing of Environment. Ann Arbor, Michigan, V.1., pp. 41-7.

Tobler, W.R., 1988. Resolution, resampling, and all that. In Mounsey, HM(ed.) Building Databases for Global Science. Taylor & Francis:London, pp.129-37.

Wolf, P.R., 1974. Elements of Photogrammetry, McGRAW-HILL:Toronto.

Système de positionnement précis et d'assiette de vol, pour la télédétection aérienne

Résumé

La recherche effectuée à l'Université de Calgary (U de C) en matière de télédétection aérienne vise notamment la mise au point d'un système de positionnement précis et d'assiette de vol, compatible avec divers types de capteurs aéroportés et qui permettra tôt ou tard l'élimination du contrôle au sol. Le présent document fait état du degré de précision nécessaire à un tel système, en précise les différentes configurations de capteurs et analyse les résultats d'essai de fonctionnement du prototype mis au point par l'université. On entrevoit pour ce système trois principaux domaines d'application, soit les applications où le positionnement précis constitue l'impératif majeur, comme en photogrammétrie, celles où les références de position et d'assiette de vol doivent être extrêmement précises, comme en imagerie multibande par capteur à barrette de détecteurs et enfin, les applications à compensation de mouvement en fonction de la vitesse, qui nécessitent une estimation précise de cette dernière, comme en radargrammétrie. Le document décrit ensuite la configuration des capteurs pour chaque groupe d'application et passe brièvement en revue les résultats d'essais exécutés au moyen de matériel aéroporté.

GPS CONTROLLED TRIANGULATION OF SINGLE FLIGHT LINES

Ayman Habib, Kurt Novak
Department of Geodetic Science and Surveying
The Ohio State University
Columbus, Ohio, 43210-1247

Abstract

Aerial triangulation controlled by GPS observations in the aircraft has been established as a precise method of photogrammetric point determination without the need of ground control. New developments of kinematic differential GPS processing yield accurate exposure locations instantaneously. If the GPS observations are available for blocks of aerial photos, the aerial triangulation can be carried out without any ground control points. Unfortunately, this method cannot be applied for single flight lines, since the GPS observations do not recover the roll angle of the aircraft. Therefore, ground control is mandatory for GPS controlled strip triangulation.

This paper investigates GPS controlled strip triangulation using known, linear features on the ground that are approximately parallel to the flight line. Point to point correspondence between the linear feature on the ground and in the image is not necessary. This described technique models the linear feature in the images by low order polynomials and forces the known line on the ground onto this function. Thus, the roll angle can be determined. In this paper, we investigate the effects of different GPS measurement accuracies both in the air and on the ground on the results. Experiments using simulated and real data are presented. We also show that this new technique is useful for mapping railroads. Our tests verify the practical feasibility of GPS controlled strip triangulation with known linear objects on the ground.

KEY WORDS: GPS controlled aerotriangulation, Strip triangulation.

1. Introduction

The implementation of the NAVSTAR Global Positioning System (GPS) changed the way in which photogrammetric images are collected and aerial blocks are controlled. There are many different applications for GPS in photogrammetry (Ackermann, 1992-a), such as:

- Precise photo flight navigation,
- Pin-point photography,
- Positioning of the perspective center for aerial triangulation, and
- GPS positioning of ground control points.

For the first and second applications real-time differential GPS based on C/A-code or P-code pseudo-range observations is required. Accuracies of 1 to 5 meters are sufficient in this case. The trajectory of the plane can be calculated continuously by observing at least four satellites in the aircraft. Pseudo range measurements are corrected by observations transmitted from a base station receiver. The computed position of the airplane can be displayed on a flight navigation system. The comparison of the airplane's actual position with the planned flight line yields corrections which are presented to the pilot or directly fed into the auto pilot system. These flight navigation systems permit the precise execution of a flight plan and the automatic exposure of aerial cameras.

The other two applications rely on more precise positioning, but do not require real-time updates. This means that both pseudo ranges and phases must be observed, in order to compute the position of the GPS antennas to within a few centimeters. The last application (control point survey with GPS) pertains to geodesy, and will not be discussed here.

The combination of GPS and photogrammetric measurements have been widely used for aerial triangulation to reduce or even eliminate ground control points. Lapine (1992) reports that aerial triangulation can be done without any ground control provided that the satellite signals are not blocked during the flight mission. However, there are a few problems that need attention in GPS aerotriangulation. Among these are the calibration of the GPS antenna offset, the camera time offset, the initial phase ambiguity, signal interruptions, and datum problems (Ackermann, 1992-b).

The GPS observations made during the photo flight in the aircraft refer to the phase center of the GPS antenna. For aerial triangulation, however, the position of the camera's perspective center is needed. The offset between the GPS antenna and the perspective center has to be determined through a calibration procedure before the flight mission (Lapine, 1990). This offset can be introduced as a constant or as stochastic apriori information in the adjustment.

Generally, the moment of exposure does not coincide with the time when the GPS receiver collects an observation. Therefore, one must interpolate between GPS positions to determine the antenna's position at the instant of exposure utilizing the time tags associated with GPS positions and the midpoint of the exposure of the photograph. Linear functions or cubic splines are commonly used for this task (Alobaida, 1993).

As mentioned earlier, GPS phase measurements are essential for achieving the accuracy required for aerotriangulation. Unfortunately, processing of GPS phase observations is complicated due to the problem of the initial ambiguity number. The integer ambiguity corresponds to the number of whole cycles the signal has traveled between emission by the satellite and its reception at the receiver. The integer ambiguity can be initialized before the flight mission from a known reference point (Lapine, 1990) or by using dual-frequency receivers and on-the-fly ambiguity resolution techniques (Schade, 1992).

If for any reason the satellite signals are interrupted during the flight, a new ambiguity number has to be found. Signal discontinuities are caused by different reasons: genuine cycle slips, interruption of the signal, and constellation changes of the satellites. Signal interruptions have been a major problem affecting GPS aerotriangulation. Different algorithms were developed for recovering the ambiguity through filtering and prediction techniques (Euler, 1990; Schade, 1992), and recently by using dual-frequency receivers.

Once GPS observations have been processed, coordinates are available in the WGS84 reference frame. Most ground coordinates, however, are defined with respect to a national coordinate system (e.g. State Plane, UTM). The transformation between these coordinate systems can be based on published formulas (Colomina, 1993) or a set of reference points available in both systems. Elevations are related to the ellipsoid and must be corrected for geoid undulations.

In order to understand the limitation of GPS controlled aerotriangulation without any ground control one can assume that the triangulation process is accomplished in two steps: relative and absolute orientations. The relative orientation can be performed by measuring at least five tie points in each stereo-pair. The resulting models can be joined together for the whole block or strip, yielding one model in a local coordinate system. Performing this task does not require any ground control. On the other hand, in order to perform the absolute orientation, control is mandatory. The minimum control requirement for the absolute orientation is three control points that must not be collinear. For GPS controlled block triangulation this condition is satisfied because the GPS observations at the perspective centers - our control - are well distributed over the whole block. On the other hand this condition is not satisfied for strip triangulation since the GPS observations of the exposure stations are almost collinear. In that case, the roll angle (around the flight line) cannot be recovered, and ground control points are necessary for solving the absolute orientation (Alobaida, 1993).

In this paper a new technique of strip triangulation is introduced that employs GPS observations at the exposure stations together with the GPS positions of linear features on the ground. In this approach, point to point correspondence along the linear feature is not necessary. This is convenient since the coordinates of the linear feature (e.g. a highway or railroad centerline) can be gathered by a moving vehicle on the ground. Thus it would be practically impossible to associate GPS coordinates with distinct physical objects along the linear feature on the ground.

Basically, the following procedure is executed:

(1) Image coordinates of a number of points are measured along the linear feature in the captured images; this can be done monoscopically.

(2) An analytical function is fit through these points in the image. Each image has an individual function representing the feature.

(3) The ground feature is projected into image space and must belong to the corresponding function in the image; this serves as a constraint in the least squares adjustment.

The next section contains a short overview of the analytics of GPS aerotriangulation. A detailed description of the newly developed strip triangulation model will be given in section 3. Results for both simulated and real data are reported in sections 4 and 5, respectively. Conclusions and recommendations are presented in section 6.

2. GPS Controlled Aerotriangulation

This section describes how GPS observations of the perspective centers can be included in bundle adjustment (Hintz and Zhao, 1989). It is assumed that GPS observations are interpolated at the exposure times of the photographs. For each camera position the observed GPS coordinates are introduced as additional observations, via equations of the form (1).

$$\begin{pmatrix} X_G \\ X_G \\ Z_G \end{pmatrix} = \begin{pmatrix} X_o \\ Y_o \\ Z_o \end{pmatrix} + \underline{R} * \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} \quad (1)$$

with:

X_G, Y_G, Z_G	observed coordinates of the GPS antenna,
X_o, Y_o, Z_o	unknown coordinates of the perspective center,
$\underline{R}$	unknown rotation matrix associated with the image,
dx, dy, dz	GPS antenna eccentricity from the perspective center with respect to the image coordinate system.

This model takes into account the eccentricity of the GPS antenna with respect to the perspective center of the aerial camera. This offset must be determined by pre-flight calibration using terrestrial techniques. We could treat these shift parameters as constants or as observations; the latter approach allows us to consider the accuracy of the calibration results. If they are treated as constants, we have regular observation equations, if they are considered observations, we have to deal with condition equations with parameters.

Additionally, linear "drift" parameters can be introduced as unknowns for each of the coordinates (Friess, 1992):

$$\begin{pmatrix} X_G \\ X_G \\ Z_G \end{pmatrix} = \begin{pmatrix} X_o \\ Y_o \\ Z_o \end{pmatrix} + \underline{R} * \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} + \begin{pmatrix} ax \\ ay \\ az \end{pmatrix} + \begin{pmatrix} bx \\ by \\ bz \end{pmatrix} * dt \quad (2)$$

with:

ax, ay, az	shift parameters,
bx, by, bz	time dependent parameters,
dt	is the time interval.

These parameters are very effective for correcting the datum transformation and cycle slips. They may be block-, strip-, or sub-strip invariant. If the GPS observations are continuous for the whole flight mission, one set of "drift" parameters is sufficient. If the satellite signal is interrupted, a new set of drift parameters has to be introduced. These parameters cause correlations and may lead to singularities. Therefore, a minimum of four ground control points and cross strips at the beginnings and ends of the strips are needed to produce accurate results (Gruen et al., 1993; Ackermann et al., 1993).

For our particular application all control is available in WGS84, both the GPS observations at the aircraft and those along the linear feature. Therefore, we do not have to worry about the datum transformation problem.

3. Strip Triangulation with a Known Linear Feature on the Ground

For this triangulation procedure we assume that GPS observations are available in the aircraft at the time of exposure. These observations are handled in the bundle adjustment according to the mathematical model shown in formula (1). As mentioned earlier, GPS controlled strip triangulation cannot be carried out without ground control, because all exposure stations are along a single flight line. Instead of ground control points, however, we utilize GPS observations along a linear feature on the ground which is approximately parallel to the flight line. These ground coordinates can be collected from a mobile mapping platform, e.g. the GPSVan (The Center For Mapping, 1991), that travels along the linear feature. The major problem of implementing features instead of points as observations for strip triangulation lies in the association of corresponding features on the ground and in the image. Since the collinearity equations are based on a point to point correspondence between image and object space, feature observations cannot be directly introduced in bundle adjustment. An extended model has to be developed to include these observations. It is explained below.

As a first step the coordinates of arbitrary points are measured along the linear feature in the image. These points are selected individually in each image. They need not to be common in successive images, because they are not used as tie points. An analytical function, usually a low order polynomial, is fit to the measured image coordinates. It represents the linear feature in that particular image. Each image has its own function which is of the form (3). The order of the polynomial can be chosen either by the operator or automatically. The polynomial coefficients and their variance-covariance matrix are determined by a least squares adjustment in a separate step for each image individually.

$$f_j(x_{l_j}, y_{l_j}, a_{o_j}, a_{1_j}, a_{2_j}, \ldots, a_{m_j}) = 0 \quad (3)$$

with:

x_{l_j}, y_{l_j}	image coordinates along the linear feature l in image j,
$a_{o_j}, \ldots\ldots\ldots, a_{m_j}$	polynomial coefficients in image j, and
m_j	order of polynomial in image j.

All GPS observations along the linear feature on the ground must belong to this function, when projected into image space. In other words, the image coordinates of theses GPS observations must satisfy (3). This constraint will be treated as an additional condition equation with parameters in the bundle adjustment. The transformation of GPS points of the linear feature into the image can be achieved with the collinearity equations (4)(5).

$$x_{l_j} = -f \frac{r_{11_j}(X_{G_j} - X_{o_j}) + r_{21_j}(Y_{G_j} - Y_{o_j}) + r_{31_j}(Z_{G_j} - Z_{o_j})}{r_{13_j}(X_{G_j} - X_{o_j}) + r_{23_j}(Y_{G_j} - Y_{o_j}) + r_{33_j}(Z_{G_j} - Z_{o_j})} \quad (4)$$

$$y_{l_j} = -f \frac{r_{12_j}(X_{G_j} - X_{o_j}) + r_{22_j}(Y_{G_j} - Y_{o_j}) + r_{32_j}(Z_{G_j} - Z_{o_j})}{r_{13_j}(X_{G_j} - X_{o_j}) + r_{23_j}(Y_{G_j} - Y_{o_j}) + r_{33_j}(Z_{G_j} - Z_{o_j})} \quad (5)$$

with:

$X_{G_j}, Y_{G_j}, Z_{G_j}$	GPS observations along the linear feature within image j,
$x_{l_j} y_{l_j}$	corresponding image coordinates,
$X_{o_j}, Y_{o_j}, Z_{o_j}$	ground coordinates of the perspective center of image j,
$r_{11_j}, \ldots\ldots\ldots, r_{33_j}$	elements of the rotation matrix of that particular image which are functions of the rotation angles $\omega_j, \phi_j, \kappa_j$, and
f	focal length.

In equation (3) the right hand sides of equations (4) and (5) are substituted for x_{l_j} and y_{l_j}, which yields (6).

$$f_j(X_{G_j}, Y_{G_j}, Z_{G_j}, X_{o_j}, Y_{o_j}, Z_{o_j}, \omega_j, \phi_j, \kappa_j, a_{o_j}, \ldots, a_{m_j}) = 0 \quad (6)$$

Thus, we introduce the orientation parameters of the photograph into this geometric constraint. Equation (6) can be treated as a condition equation with parameters in the bundle adjustment. In this equation, the perspective center coordinates and rotation angles are unknowns, while the GPS observations along the linear feature and the polynomial coefficients are observed quantities. This model also accounts for the accuracy of the observed GPS coordinates along the linear feature and the computed polynomial coefficients as obtained from (3).

There are two practical problems in implementing this algorithm. Namely, choosing the GPS observations along the linear feature that belong to a particular image and selecting the order of the polynomial.

The GPS observations along the linear feature within image j can be obtained as follows: the corners of the image are projected into object space using the initial values of the exterior orientation parameters of image j as obtained from the GPS observation in the aircraft and the approximations of the rotation angles. The GPS observations along the linear feature that fall within the projected area plus a threshold are selected.

The order of the polynomial can be chosen fully automatically by statistical testing of the computed parameters. We begin with a high order polynomial and check the significance of each coefficient by testing of hypothesis, (7), which tests whether the coefficient under consideration is significantly different from zero.

$$H_o: a_{i_j} \sim N(0, \sigma_i^2) \qquad i = 2, \ldots m \qquad (7)$$

with

σ_i^2	is the variance of coefficient a_{i_j} as obtained from equation (3), and
a_{i_j}	highest order coefficient of the polynomial.

This null hypothesis (H_o) can be tested through Chi-squared tables, as follows:

if $a_i^2\sigma_i^{-2} \leq \chi^2(\alpha, df)$ accept H_o

otherwise, reject H_o

where:

α is the chosen level of significance, and

df is the degree of freedom, 1 in this case.

If the null hypothesis is accepted, which means that the specific coefficient equals 0 and does not contribute to the form of the curve, the polynomial order is reduced by one. This test is carried out for each coefficient from higher to lower orders until the null hypothesis is rejected.

4. Results with Simulated Data

To evaluate the performance of the new model, many experiments were conducted using simulated data. We assumed a strip of four images containing one linear feature. The interior orientation parameters of the camera and the exterior orientation parameters of the images constituting the strip are pre-defined. Along these images both tie points and feature points were simulated. The image coordinates of all object points and features were computed using the known exterior orientation of each photograph.

Experiments were conducted using the simulated image coordinates of the tie points and points along the linear feature as image observations (degraded by random errors of image coordinate measurement), and the exposure station coordinates and the linear ground feature as GPS measurements. We wanted to find out how the new technique behaves under different conditions, such as varying GPS accuracies at the perspective centers and on the ground. Its performance was evaluated based on the differences between the estimated ground coordinates of the tie points and their real values. In figures 1 to 4, these differences are displayed as error vectors originating from the true position of each point to the computed one in X and Y directions respectively.

Figure 1 shows the results obtained by assuming a GPS accuracy of 1.0 m both on the ground and at the exposures stations. The mean error at the tie points after triangulation is about 0.5 m, which is quite acceptable considering the low accuracy of the GPS observations. A closer investigating of figure 1 reveals that these deviations correspond to a scale error. In other words, the accurate spacing between the exposures (base length) were not completely recovered along the flight line. This is expected since the linear feature is almost parallel to the flight direction, which makes it ineffective for recovering this component. It only solves for the roll angle, while the shift, the scale, the direction, and the pitch are solely recovered by the GPS observations of the perspective centers.

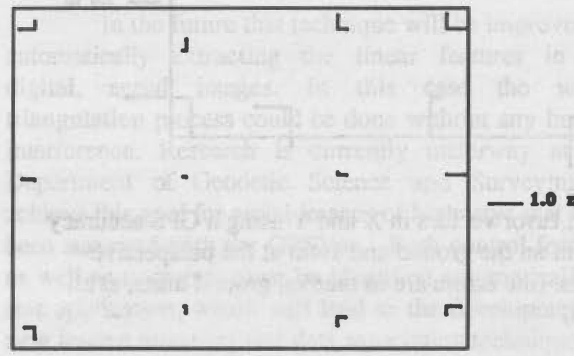


Fig. 1: Error vectors in X and Y using a GPS accuracy of 1.0m on the ground and 1.0m at the perspective centers. (the errors are in meters, ground units, at all points)

Figure 2 shows the results obtained by assuming GPS accuracies of 1.7 m and 1.4 m on the ground and at the exposure stations respectively. Figure 3 presents the errors at the computed tie points under the same conditions but with unknown interior orientation parameters. It is clear from this figure that the system is not capable of solving for the x-coordinate of the principal point due to its high correlation with the X-component of the exterior orientation parameters. Figure 4 shows the results obtained by assuming a GPS accuracy of 1.7 m on the ground and 2.0 m at the perspective centers. This figure indicates an ill-conditioning in the normal equation matrix (i.e. the results are not reliable any more).

Table 1 shows the rms values in X, Y, and Z direction for the experiments described by Figs 1 to 4:

GPS Accuracy (m)		RMS (m)		
Air	Ground	X	Y	Z
1.0	1.0	0.37	0.28	0.45
1.4	1.7	1.67	1.24	1.86
1.4	1.7 (sc)	3.95	0.45	6.98
2.0	1.7	13.42	9.94	14.49

(sc) self calibration of the camera.

Table (1), Rms values in X, Y, and Z directions for different GPS accuracies both at the perspective centers and along the linear feature on the ground.

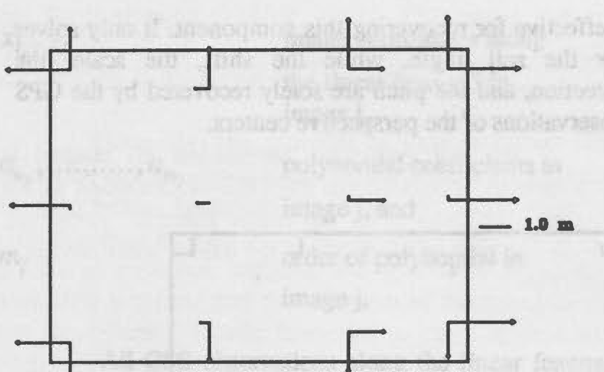


Fig. 2: Error vectors in X and Y using a GPS accuracy of 1.7m on the ground and 1.4m at the perspective centers. (the errors are in meters, ground units, at all points)

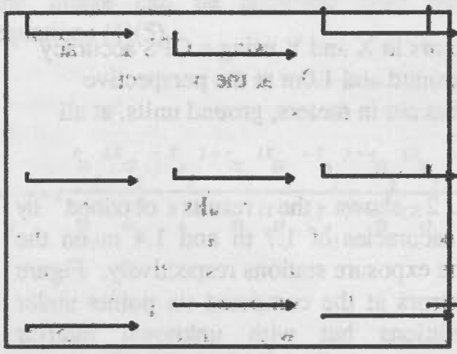


Fig. 3: Error vectors in X and Y using a GPS accuracy of 1.7m on the ground and 1.4m at the perspective centers (the errors are in meters, ground units, at all points) with unknown interior orientation parameters.

Fig. 4: Error vectors in X and Y using a GPS accuracy of 1.7m on the ground and 2.0m at the perspective centers. (the errors are in meters, ground units, at all points)

In summary, the linear feature only overcomes the datum deficiency due to the inability of solving for the roll angle. Thus, it will contribute nothing towards computing the other orientation parameters.

5. Results with Real Data

This technique was applied to triangulate strips of digital aerial photos captured along a railroad line. This project was conducted by the Center for Mapping of The Ohio State University together with Ruekert and Mielke (Waukesha, WI). It was sponsored by Conrail (Philadelphia, PA) and MCI (Houston, TX). GPS observations were captured along the rail road by loading the Center's GPSVan on a flat car, which was pulled by a locomotive for 50 miles (figure 5). The van collected both GPS positions to define the rail centerlines, as well as digital image pairs and analog videos. Digital aerial images were collected by The Sewall Company (Old Town, Maine) using a Kodak DCS camera (1280x1024 pixel frame CCD sensor) and a Trimble 4000ST GPS receiver. The results of GPS controlled strip triangulation (tie point coordinates) were used to rectify the aerial images to generate digital orthophotos of the railroad tracks and the surrounding areas.

In this project the GPS positions at the aircraft were computed with a standard deviation of 3.5 m. The rail-centerlines are accurate to about 1.5 m. Image coordinates of 10 to 15 points per model and 5 to 6 points along the linear feature were measured monoscopically on a computer monitor. We estimate that the measurement accuracy is better than 1 pixel which corresponds to 30 cm on the ground. Using these image coordinates the polynomials were interpolated and the bundle adjustment was computed. The variance component was typically two to three pixels large. All together 21 strips of about 8 photos each were triangulated.

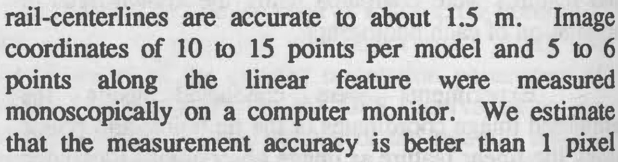


Fig. 5: GPSVan loaded on a flat-car, Stereo-vision system.

The resulting product of this process is shown in figure 6: the rail centerlines captured by the GPSVan are super-imposed on top of rectified and stitched aerial images.

6. Conclusions and Recommendations

A new technique for GPS controlled strip triangulation using linear features as control was presented. It was tested using both simulated and real data. The obtained results verify the feasibility of the proposed method. The method is based on the collection of ground coordinates along a linear object using a mobile mapping system, such as the Center for Mapping's GPSVan. As mobile mapping systems can also be equipped with a stereo vision system for spatial feature collection from terrestrial image pairs, it is possible to identify a few ground control points in each strip. This is of advantage, as it improves the accuracy of the results, especially if the interior orientation of the camera is not known precisely. Other applications for GPS controlled strip triangulation include the mapping of utility lines, pipelines, and highways.

In the future this technique will be improved by automatically extracting the linear features in the digital, aerial images. In this case the whole triangulation process could be done without any human interference. Research is currently underway at the Department of Geodetic Science and Surveying to achieve this goal for aerial images of highways that have been surveyed with the GPSVan. Both control features as well as tie points must be identified automatically in this application, which will lead to the development of new feature matching and data association techniques.

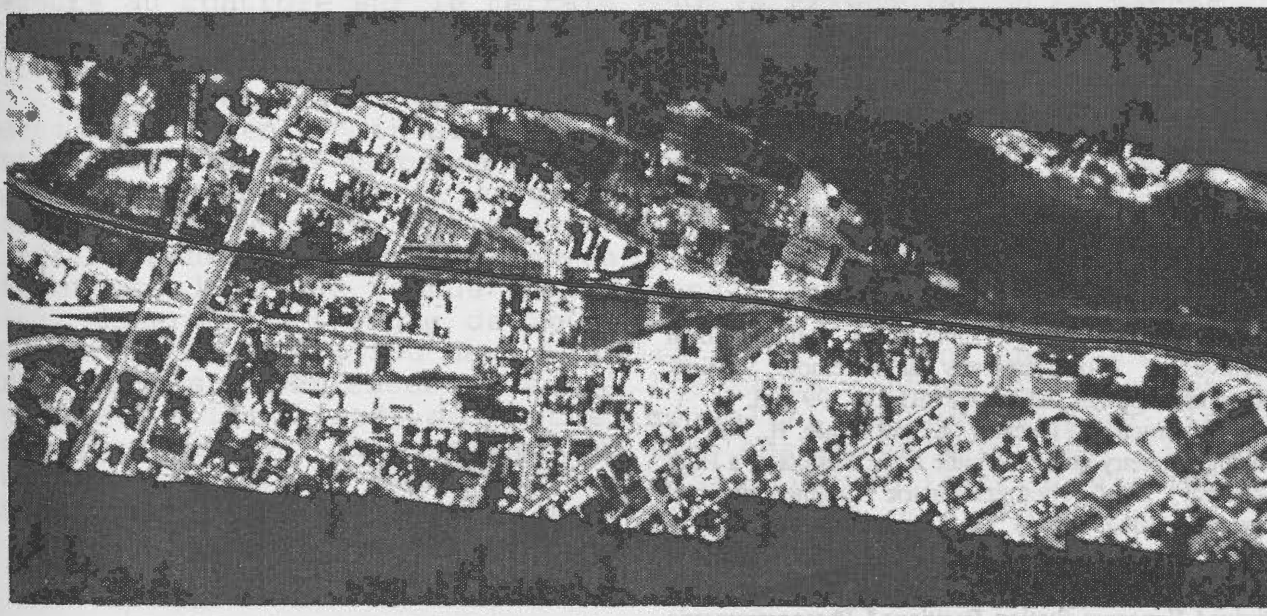


Fig. 6: Rectified and stitched digital aerial images overlaid by rail road center-line.

REFERENCES:

Ackermann, F., 1992-a. "Kinematic GPS Control for Photogrammetry." Photogrammetric Record, Vol. 14, No. 80, pp. 261-276.

Ackermann, F., 1992-b. "Operational Rules and Accuracy Models for GPS-Aerotriangulation." International Archives of Photogrammetry and Remote Sensing, Commission III, Vol. XXIX, Part B3, pp. 691-700.

Ackermann, F., and Schade, H., 1993. "Application of GPS for Aerial Triangulation." Photogrammetric Engineering and Remote Sensing, Vol. 59, No. 11, pp. 1625-1632.

Alobaida, A., 1993. "Design and Simulation of a Real-Time Mapping Satellite for the Kingdom of Saudi Arabia." Dept. of Geodetic Science and Surveying, The Ohio State University, Ph. D. Dissertation.

Colomina, I., 1993. "A Note on the Analytics of Aerial Triangulation with GPS Aerial Control.", Photogrammetric Engineering and Remote Sensing, Vol. 59, No. 11, pp. 1619-1624.

Euler, H., Goad, C., 1991. "On Optimal Filtering of GPS Dual Frequency Observations without Using Orbit Information." Bulletin Geodesique, August, Vol.65, No. 2, pp. 130-143.

Friess, P., Heuchel, T., 1992. "Experience with GPS-Supported Aerial Triangulation." International Archives of Photogrammetry and Remote Sensing, Commission I, Vol. XXIX, Part B1, pp. 299-305.

Gruen, A., Cocard, M., and Kahle, H., 1993. " Photogrammetry and Kinematic GPS: Results of a High Accuracy Test." Photogrammetric Engineering and Remote Sensing, Vol. 59, No. 11, pp. 1643-1650.

Hintz, R., and Zhao, M., 1989. "Considerations in the implementation of Aerotriangulation with GPS Derived Exposure Station Positions." Photogrammetric Engineering and Remote Sensing, Vol. 55, No. 12, pp. 1731-1735.

Lapine, L., 1990. "Analytical Calibration of the Airborne Photogrammetric System Using A Priori Knowledge of the Exposure Station Obtained from Global Positioning System Technique." Dept. of Geodetic Science and Surveying, The Ohio State University, UMI Dissertation Services, Publication No. 9111738, Ann Arbor, MI 48109.

Schade, H., 1992. "Reduction of Systematic Errors in GPS-Based Photogrammetry by Fast Ambiguity Resolution Techniques." International Archives of Photogrammetry and Remote Sensing, Commission I, Vol. XXIX, Part B1, pp. 223-228.

The Center for Mapping, 1991. "GPS/Imaging/GIS Project." Final Report, Center for Mapping, The Ohio State University.

TRIANGULATION D'AXES DE PASSAGE UNIQUES À L'AIDE DU SPG

Résumé

La triangulation aérienne au moyen d'observations SPG à partir de l'aéronef a été établie comme une méthode précise de détermination photogrammétrique des points, sans contrôle sur le terrain. Les derniers progrès dans le traitement SPG cinématique permettent de situer précisément et instantanément les expositions. Il est possible d'effectuer une triangulation aérienne sans aucun point de contrôle sur le terrain, à condition de disposer d'observations SPG pour des blocs de photographies aériennes. On ne peut toutefois appliquer cette méthode à des axes de passages uniques car les observations SPG n'incluent pas l'angle de roulis de l'aéronef. Il est donc essentiel d'avoir recours au contrôle sur le terrain pour la triangulation de bandes à l'aide du SPG.

Cet article étudie la technique de triangulation de bandes, à partir d'éléments linéaires terrestres connus et à peu près parallèles à l'axe de passage. Il n'est pas nécessaire d'établir la correspondance entre les points situés le long de l'élément linéaire et ceux de l'image. Le système modélise l'élément linéaire dans les images au moyen de polynômes d'ordre inférieur et intègre la ligne connue au sol dans cette fonction. On peut ainsi déterminer l'angle de roulis. Les effets des différentes précisions de mesure SPG dans les airs et au sol sont analysés. On présente les expériences menées à l'aide de données simulées et réelles. L'utilité de cette nouvelle technique pour la cartographie des chemins de fer est également montrée. Les résultats des essais confirment la faisabilité de la triangulation de bandes à l'aide du SPG pour des objets linéaires connus au sol.

LARGE SPATIAL OBJECT HANDLING IN GEOGRAPHIC INFORMATION SYSTEMS

Wenjin Zhou
UNISYS SYSTEM 9 GIS
61 Middlefield Road
Scarborough, Ontario, Canada M1S 5A9

ABSTRACT

We have seen the increasing availability of large spatial objects in Geographic Information Systems(GISs) in the past two decades, mainly due to the interest of multi-media information systems. These large objects may come from remotely sensed imagery, scanned maps, digital terrain models, and simulation model output. Due to the large amount of data and rich semantics involved in the large objects, traditional methods used to handle text and graphics data will not work well for large spatial objects. Therefore how to effectively and efficiently store, manage and process large spatial object presents a special challenge to the current generation of GISs.

This paper presents techniques that can be used to handle large spatial objects. After analyzing the nature of large spatial objects, we decompose them into three components: row data, Direct Related Attribute Data and Derived Attribute Data in order to make them more manageable in GISs. We then investigate GIS query patterns on large spatial objects and present an object oriented data model that uses objects, methods, rules, and class browser to handle large spatial objects in the context of GISs. We also present five abstraction techniques that can be implemented to effectively and efficiently support large objects in GISs. Finally we use several practical examples to demonstrate the effectiveness of the suggested methods. We conclude the paper with a short discussion of system performance and further research issues.

KEY WORDS: Large spatial object, Object oriented model, Data base, Query Patterns, Abstraction, GIS

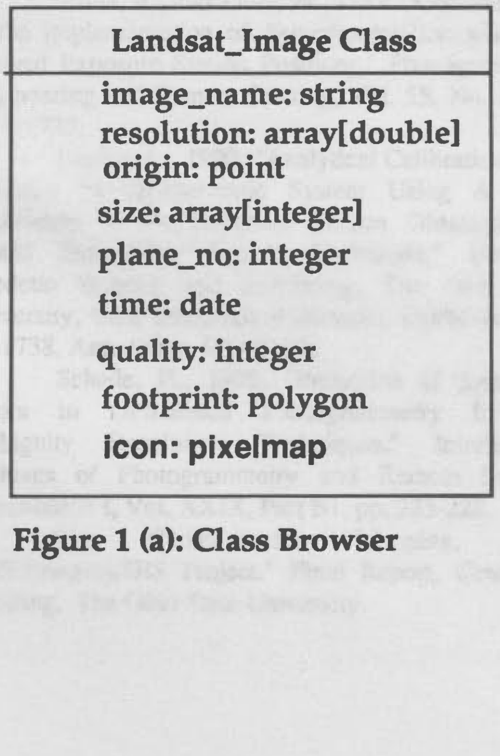


Figure 1 (a): Class Browser

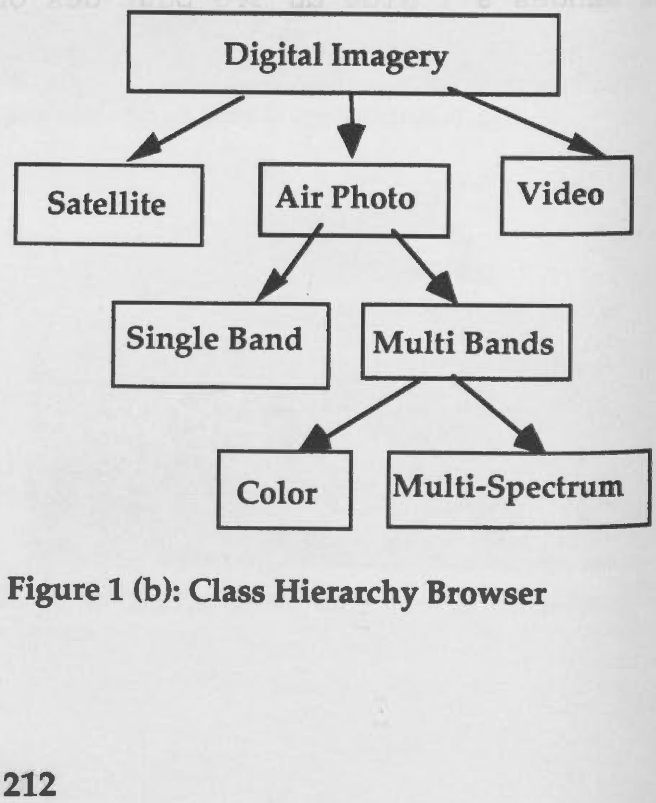


Figure 1 (b): Class Hierarchy Browser

1. INTRODUCTION

In the past few years, a substantial development has been going on in the field of managing large spatial objects such as digital imagery, digital terrain model and scanned maps, mainly due to the interest of building multi-media spatial information systems and global environmental information systems. The development can be roughly classified into two categories. The first category is through system integration which uses two or more different systems, such as image processing system to handle image objects and Data Base Management System(DBMS) to handle non image objects such as text and graphics. Examples can be found in Chang[1990], Zhou Q.[1989], Wegener[1989], and Zhou[1991]]. The second category is through next generation data base systems using Abstract Data Type(ADT) or object oriented data model to handle large objects. These systems include Lohman[1989], Orenstein[1989], Deux [1990], Gupta[1992], and Stonebraker [1993].

However, there are no generally accepted solutions in GISs at this time. With the first method, two systems are loosely integrated. Large objects in the image processing systems are processed independently and the results are converted into the DBMS to perform GIS operations. This not only limits the use of DBMS for large object management, but also makes the data processing unnecessary complicated and time consuming because of multi data conversions. With the second method, all large objects are treated as long binary data strings with little semantics and data abstraction associated with them. This is not only inefficient for data processing because the whole data set may need to be read, written and processed together, but it also makes many kinds of interactive data processing impossible.

In this paper we will analyze large object contents, highlight their special features, and develop an object oriented model to support them in GISs, using digital images as examples. We will also investigate their query patterns and present several methods that can be used to reduce the amount of data, to improve data retrieval efficiency, to speed up data query and to better support browsing. We will then use several practical GIS query examples to show the performance improvement upon using different techniques. We conclude the paper with some discussions on system performance and future research issues.

2. LARGE SPATIAL OBJECTS

Large spatial objects are often represented by multi-dimensional matrices using long unstructured byte strings that are often stored and transmitted entirely. More precisely large objects consist of a list of small items and long data strings. The list of small items will be used to interpret the data format and meaning of the unformatted long data strings following it. For image data, these small items may be the image header; For Digital Terrein Model(DTM) data, these may be the name of the region, the coordinates of the origin, resolution, precision, etc. These small items are often mandatory and are used to interpret long data string for display and process, and/or to identify and distinguish one data string from others. We call these small items *direct related attribute data(DRAD)*. While DRAD is indispensable, other formatted attribute data describing the contents and features of large spatial objects is generally not mandatory. For digital imagery, this data may be histogram, color map, and interpretation results from the original image data; For DTM, it may be the contour line, the slope and visibility data. This data is often the result of data calibration, interpretation, processing and analyses. We call this data *derived attribute data(DAD)*. The DAD data is per se redundant because it is just another form of information presented in the source data. Usually DAD is very difficult and/or time consuming to derive. In GIS, it is desirable to store DAD in the database because DAD is high level information and can be used to answer most GIS queries. In addition we prefer to integrate DAD into GIS databases because DBMS then can be used to manage DAD.

Because DAD and DRAD is much simpler than the raw data and sometime they may be well modeled by relational data model, many researchers use this technique to handle large objects. They store DAD and DRAD in RDMS, while manipulating the long string data through a link between the relational table and operating system files[Zhou, 1991]. This approach may work for simple applications, but it has several serious drawbacks. First, the DAD and DRAD can be semantic rich and complicated (for example geometry and topology data) because of the amount of information embodied in a large object. Second, because of two data bases used are independent, it is very difficult to maintain data integrity and perform transaction management. Third, the relational data model lacks the power to define the semantics inherited in the large objects and the methods needed to process large objects. It is indeed not too much difficult to retrieve an image by its name or number(or other identifiers), but it may be for more difficult to do it through selection criteria that would be contained in the image. This may lead to very long research, or even sometimes to tedious image processing (e.g. for similarity retrieval). This observation leads us to consider a more powerful data model and processing methods that can be used to handle large objects.

3. LARGE SPATIAL OBJECT QUERY PATTERNS

Before we discuss new data model and processing methods that can be used to improve the system performance on large objects handling, we need to examine how users will make use of large spatial objects in GISs. First, many queries are concerned about the general information of large objects and these can be classified as requests of meta-data, i.e. information about large objects. The common characteristics of this type of queries are that queries touch a high volume of data, but the generated answers are tiny by comparison. With our model, it is not difficult to figure out that this type of queries can be easily answered using DAD or DRAD.

The second type of query is a kind of data mining. Users try to query over a large set of objects in order to find a small set of specific data. For example, a user may want to find an image that has a certain percentage of the snow coverage. This type of queries does not require the exact and detailed information, but instead requires the "overview" of the special nature of a large object. In this case, low resolution images with certain rules specified in a query language to derive the required results are possible solution for this purpose.

The third type of query displays a large spatial object on a computer screen in order to do image interpretation or image analyses. Due to the small size of computer screen, only a very tiny part of the large object can be displayed on the screen at a time with the original data resolution. In this case we can significantly improve the system performance by windowing the image first and then only render the small part of data on the screen. By buffering the nearby part of the image, we can also perform zoom/pan in a very fast fashion. This can be very significant when we compare to the case where the whole image data needs to be read from a data base.

The fourth type of query is data browsing. Users wish to perform data browsing because they are not sure which type of data they are dealing with. Class hierarchy or data base schema browsing can be very helpful to the new users. Useful summary information includes histograms showing data statistical feature, images outline showing the data extent and well-designed icons showing data characteristics. For GIS applications it is our belief that all these techniques are critical to the success of GISs when they are used to handle large spatial objects.

4. OBJECT ORIENTED DATA MODEL

Before we present our new data model, let us consider a simple example of a large image object. The image is a Landsat image with DRAD such as image dimension, the geographical location of the image origin, the band number, the pixel precision, the acquisition time, the control points used for image geometric rectification, the parameters used for radiation correction, the data quality, etc.. The DAD data for the Landsat image may be its histogram, classification result(a set of polygons with category numbers), a two dimensional array for the classification precision, a set of rules to derive a certain theme (for example, a rule like: "if a pixel value is large than 10 and less than 20, then the pixel can be classified as water".) It is difficult if not possible for the relational data model to handle this non primitive data (e.g. point, polygon, array), rules and data operation methods. However, they can be well handle by an object oriented data model [Zhou and Wilkinson, 1993]. In the above example, we could define an object class Landsat_image, with various attributes to store meta-data, with methods to present the correction models, and with class rules to present the data classification rules[Hughes, 1993].

Two most often used methods to handle large objects in object oriented models are: first, the extended relational DBMSs use Abstract Data Types(ADTs) to support large and unstructured data such as images[Stonebraker, 1993]. It allows users to define their own functions to manipulate the ADTs. This method allows the system to be extended, but leaves the extensions to its users. In addition, the raw data is treated as unformatted byte strings and very little semantic information is maintained in the system. The second method stores large objects in user defined classes, and allows the data base administrator to define methods and other attributes for classes[Deux et al., 1991]. Usually both methods store raw data using Binary Large OBjects(BLOB), and provide the capability to store and retrieve them through a query language.

In both methods the row data is treated as a whole and no abstraction methods are provided to the users. In addition most commercial systems generally do not allow the query language to be extended to capture the semantics of large objects. Therefore new functions and operators defined by users may not be able to use internal data indexing and clustering techniques provided by the system to manage BLOBs and make the user defined methods or operators inefficient.

To make large object manageable within a GIS environment, some new methods such as data compression, data ordering, controlled data redundancy, indexing and data browsing needs to be investigated, evaluated and implemented based on our data model and the possible GIS query patterns. This paper will mainly discuss these methods and present our implementation strategies.

Using the object oriented data model, we have implemented a visual based schema browser to help users to define and modify the data base schema[Zhou and Wilkinson, 1993]. An example of the data base schema browser is shown in Figure 1.

5. LARGE SPATIAL OBJECT HANDLING METHODS

Many current generation GISs have been designed to effectively manage point, line, surface and their attribute data. However, large spatial objects such as remotely sensed image, scanned document and DTM are series of large multidimensional arrays with various temporal, spatial and spectral characteristics. They are unformatted and quite often require long delayed processing in order to answer user's query or provide other useful information. Interactive, network based GIS queries may be very slow when large objects are concerned. Therefore efficient support of large objects must be designed into GISs.

In order to manage large spatial objects effectively and efficiently, special techniques that allow applications to quickly store and retrieve large spatial object onto and from secondary memory devices such as a disk need to be developed. The traditional methods of storing large objects, such as imagery is linear allocation where the data is laid out linearly by a nested traversal of very dimension in a predefined order. The method may be optimized when the whole data of the large object is needed. Due to the limitation of the current viewing system such as computer screen(e.g. 1000 by 1000 screen pixels vs. 20 by 20 K scanned map) and the nature of data usage(e.g. quite often only a small part of data is used at a time), the linear type data storage can be very inefficient in many GIS applications. This situation will get worse when the dimension and size of data increase.

5.1 Data Compression

One area that needs to be investigated is how we can reduce the amount of large object data while keeping most GIS queries satisfied. One obvious answer is data compression techniques such as reducing data quantization level or resolution, or using Run Length Encoding and JPEG compression [Wallace, 1991]. These techniques can significantly reduce the size of large objects and make large objects more manageable in data bases.

The compression is one of the techniques for data reduction. It takes advantage of patterns in image data by removing redundancies and in effect, squeezing the data to maximize the information contained in each byte. Scanned maps, digital air photos and satellite images are considered good candidates for compression because of high correlation in the image data. The compression can be lossless(e.g. the decompressed data is identical to the original, Run Length compression is an example) or lossy (some information is lost. JPEG is an example for this type of compression). Lossy compression generally yields higher compression ratios than lossless. For some GIS applications, such as image display, this is a reasonable trade-off. However, some applications will not allow the lossy compression because of the precision requirement.

Depending on the nature of image data, either lossless or lossy compression may significantly reduce the size of the image data to be managed. For scanned maps, lossless compression typically reduces the size by a factor of 2-10. This allows the database system to transfer more information with fewer bytes, thus making effective use of available band width. This can be very significant for network based GIS systems.

Another method to perform data reduction is to reduce the data resolution and quantization level. This technique can significantly reduce the amount of data while satisfying most GIS queries. Table 1 gives the example of various data reduction methods and their data sizes before and after the data reduction. Since most GIS queries are "Gold Mining" or "Landscape" type that do not require detailed information in the first several queries, this technique can be significant for GIS applications.

5.2 Ordering

The second area that we want to look into is to explore the internal nature of the large spatial objects. For example, it has been demonstrated that remotely sensed imagery has very high spatial correlation among pixels in a neighbor and between various spectral bands. Partition and re-ordering the spatial data to take these features into consideration can speed up many GIS queries.

The traditional method of storing large object data is linear allocation whereby the data is stored by some predetermined orders. Typical date storage orders may be one of the pixel interleave, line interleave or plane interleave. This order may be efficient for some data processing that involved all pixels, but will make other access such as point or window queries in-efficient. Optimizing the allocation of large object data becomes increasingly important as more and more multi-spectral, multi-temporal large objects are available today in GISs.

In GIS application such as display, editing and analysis, only a small part of data is required at a time. Point or window based queries to extract parts of the image are most often performed. This makes data ordering particular important for large object handling. To improve the system performance, several strategies have been proposed, which include:

a) Tiling: which divides an image into tiles that are stored and accessed independently. By combination with parameters such as optimism buffer size, display window size, these techniques can improve the system performance dramatically, especially for network based GISs.

b) Reordering: which permutes the order of image pixels to reduce average seek distance. For example, in data display, we could anticipate that zoom/pan is frequently performed on the large objects. Therefore, instead of linear storing each tile, we use the Mount-Caro order[Samet, 1989] to organize the tile files.

c) Redundancy: which stores redundant copies of the image. The redundant images are abstracted and organized differently to optimize different patterns of access. For example, for overview purpose, a low resolution image can be used to give users the rough idea on the object. Only when the user specifically zooms in the data, the original data is provided to give more details. This technique not only makes the system easy to use, but also makes most interactive applications possible.

5.3 Abstraction

The third area that needs to be investigated is object abstraction. The conventional DBMS paradigm handles well-formatted data and does not need to perform any abstraction on the data. In contrast, large spatial objects are intrinsically information intensive and semantics richer. Intelligence needs to be build into the system before the system can be used to effectively and efficiently answer user's queries. The DAD in our data model is mainly used for this purpose, because we believe that data abstraction can be built into DAD either through automatic image processing methods or through manually editing or indexing methods and can be used to answer many queries in GISs.

Abstraction can be defined as the process of constructing new concepts from a given collection of concepts. Abstraction is a method of extraction of the essential part of the data set. It is created using some combination of statistical analyses, data classification and artificial intelligence. Derived Attribute Data is the result of abstraction. Classification in the object oriented modeling is the example of abstraction. Image classification is another example of abstraction.

The key observation is that while a query may be posed to a large object, the answer to the query may be tiny. For example, a query such as "what is the size of Lake Ontario" can be answered very easily if the abstracted DAD data is available. While many techniques can be used to perform data abstraction[Sequio 2000 Report 11], we only limit ourselves to the construction of meta-data and the vector counterpart of large spatial objects. For example, we use statistical analyses to derive the mean, variance, histogram, and other parameters of image objects. We also use image classification techniques and rules to extract classes such as water, cloud, land use, etc. with manual editing techniques followed to derive vector representation of the image. These derived data is modeled as DAD data and is stored in GIS data base for various applications.

5.4 Indexing

The last method to be investigated is ability to provide value indexing. The conventional DBMS paradigm provides object indexing. Hence, one can designate one or more fields in a record as indexed, and DBMS will build the appropriate kind of index on the data in the required fields. Value indexing may be reasonable in traditional applications, but will not work for the type of large objects such as remotely sensed data. First, queries are issued to retrieve images by their content, e.g. to find all images that have Lake Ontario. This search requires indexes on the result of a classification function (on virtual attributes) and not on raw images. Second, indexing functions for images and text often return a collection of values for which efficient access is desired. For example, a keyword extraction function might return a set of relevant keywords for a document, and the user desires indexing on all keywords.

To be able to retrieve large object stored in databases we must associate identifying parameters with each of the large objects. It is through these parameters that we can select stored large objects for display, comparison, queries and further analysis. Primary parameters are produced when large objects are obtained and describe the large object's acquisition procedures and it's related properties. For fast systematic analyses and access we need to retrieve images based upon their contents. The pixels that make the image are rarely suitable for direct search. If we can provide secondary parameters indicating the location and type of objects seen on the image, the database searches could be enabled that directly serve the end-user' objectives. Modern GISs without such a retrieval capability will not be justifiable.

However, the majority of secondary parameters used today, when available at all, is entered by humans after visually scanning an image [Wiederhol 1989 and Chang 1992], This observation motivates us to use abstraction techniques to generate the secondary parameters. Although a completely automated abstraction is our long term goal, only semi-automatic analysis and classification are used in this research to generate the DAD that is used as secondary parameters in the content based queries.

6. PERFORMANCE EVALUATION

Before we look into our examples, let us first consider how big some typical spatial objects may be in GISs:

a) a scanned 8 bits map data 10, 000 by 10,000 pixels without compression is about 100 MB;

b) a SPOT panchromatic image 6000 by 6000 is about 36 MB;

c) a Landsat Thematic frame(7 bands) is about 300 MB;

With more than 1000 large objects in a GIS database, we can imagine how important it is to effectively and efficiently store, transfer and process these large objects. The methods used to handle these large spatial objects will mainly decide the success of GISs.

6.1 Application Examples

Without practical examples with real GIS queries to large spatial objects, we can only speculate on the efficiency of the methods we presented. In the previous section, we have analyzed types of queries a user may actually run against large spatial objects and shown how often the abstracted data can be used to answer user's queries. In this section we will demonstrate the difference in system performance when our approaches are used.

To show the performance improvement using our proposed methods, we use a data base that consists of ten large image objects(about 50 MB), 5 MB vector and attribute data(including DRAD and DAD). We consider this database very small compared to real GIS data bases. This database is stored in the disk on a SUN SPARC 2 station with 32 MB memory. We post the following five different types of queries on large spatial objects in the database:

1) Overview: display an image in a window;
2) Zoom/pan: Zoom/pan an image.
3) Thematic selections: Show me the percentage of "Water Coverage" in a large image object.
4) Range Query: Show the image that is contained in a polygon ABCD.
5) Automatic analysis: Detect all edges in a Landsat image using a predefined method"

These selections are often followed by further processing steps in practical applications to refine the selections. This is an important feature in large object queries. We will not discuss this in this paper because this refinement is very specific for each particular application.

6.2 Query Cost Comparisons

a) Query One: Image Overview

Using the traditional method, we would need to read all image data from the database into memory, zoom out the image to make it fit in the window and then display image data on the window. This will take several minutes to complete.

In the proposed method, we just retrieve the overview image that is generated using lower data quantization and spatial resolution at the time when the large image is loaded into the data base. This would reduce the time of the image overview into seconds.

Table 2 gives the result when the query is applied to a 2000 by 2000 pixels 24-bit image and an 4-bit overview image of 512 by 512. The overview image is generated by an edge preserved reduction technique. From this result, it has shown about 10 times time saving.

b) Query Two: Zoom/pan an Image

Using Binary Large Object without image tilling and reordering, we need to seek through the whole image data, read the needed part of data, and then transform the image data using a required zoom/pan parameter.

With proposed techniques such as tilling and data reordering, we tile the image first and reordering the tiles according to their spatial relationship. In case of zoom/pan, we always load 8 nearest tiles of the current tile into memory and then zoom/pan is a matter of deciding which pixels are "visible" for current query.

Using an 1-bit image with size of 8000 by 8000 pixels and a tile size of 512 by 512, the query cost using new method is shown in Table 2 and is about 30 times fast.

c) Query Three: Thematic Selection

Without any DAD stored in the database, we need to retrieve the image using the image name. Then we read all image data into memory and use a rule to derive the "water coverage." Last, we calculate the size of the "water coverage" by counting the total pixel numbers. This is a very slow and tedious query processing.

With our proposed method, we derive all thematic data that include "cloud coverage" theme using image classification and store them as DAD in the data base at the time when the image is loaded into the database. In this case, we store the theme "cloud coverage" as a polygon map with attributes associated with each individual polygon. When a thematic query is issued, it is just a matter of retrieving the right polygon from the database.

Table 2 gives the result when the query is posed on our test database against a satellite image with the size of 1024 by 1024 pixels and the quantization level of 24 bits. From the table, the time difference between two methods is 2:15.

d) Query Four: Range Query

Using Binary Large Object without any data abstraction, we need to do the following three steps: first, perform spatial searching to get all image objects in the polygon ABCD; second, retrieve all image objects into memory using the image ids; third clip each image object against the polygon and return the clipped data to the application.

In the proposed method, we tile all image files into small files and build a Minimum Enclosing Rectangle of each tile into DAD in the database. Then any range query is just a query that retrieves those tiles that intersect with the polygon ABCD and mosaic them together. This not only significantly reduces the amount of data involved, but also improves the processing speed.

Table 2 shows a result using a scanned 1-bit map with size of 8000 by 8000 and a polygon size of 256 by 256 pixels. It can be seen that the query cost is 1: 81.

e) Query Five: Automatic Analysis

Without abstraction, we need to execute a function that can detect the existence of edges on the original image and display the result. This may take several minutes to finish.

Instead of doing this we execute the function against the abstracted image(4 times resolution reduction) in the database. We can get almost the same result, but with significant performance improvement.

Using a 24-bit 2000 by 2000 scanned map as an example, the query cost using two different methods is shown in Table 2. It is easy to see that the suggested method is about 15 times fast than the old method.

From these examples we have concluded that the suggested techniques can improve the GIS query on large objects significantly. The initial cost to set up the DAD and DRAD in GIS data base is relatively small when we consider that GIS database can be used by multi users at various times.

7. CONCLUSIONS AND RECOMMENDATIONS

This research has shown that various techniques can be implemented in GISs to improve system performance on large spatial objects handling. Our data model distinguishes DAD, DRAD and Row Data in large spatial objects and is implemented using an object oriented data model. We discuss the query's patterns of large objects in GIS applications and propose an integrated technique that uses data compression, ordering, abstraction, indexing and browsing to improve the system performance on large object handling. Using several examples we have shown that these techniques can significantly improve the system performance, especially for network based GISs.

The future work includes developing methods that can automatically generate and update DRAD and DAD in GIS databases. We plan to use export system in this respect. In addition, we plan to do more tests on the performance of the network based GISs using the proposed methods.

ACKNOWLEDGMENTS

Special thanks go to Steve Wilkinson for going through initial drafts of the paper and providing useful feedback and UNISYS GIS for providing the time to work on the paper.

REFERENCES

Chang S.[1990]: "Principles of Pictorial Information System Design." Prentice Hall.

Chang S. and A. Hsu[1992]: "Image Information Systems: Where Do We Go From Here?" IEEE Transactions On Knowledge and Data Engineering. Vol. 4, No. 5, October.

Deux O. et al.[1991]: "The O2 System.". Communications of The ACM. Vol. 34, No. 10, October.

Gupta A., T. Weymouth and R. Jain[1992]: "Semantic Queries in Image Databases." Visual Database Systems, II E. Knuth and L.M. Wegner(Editors) Elsevier Science Publishers B. V. (North-Holland).

Hughes, J. G.[1991]: "Object-oriented Databases." Prentice Hall International.

Lohman G., B. Lindsay, H. Priahesh and K. Schiefer[1991]: "Extensions to Starburst: Objects, Types, Functions and Rules." Communications of The ACM. Vol. 34, No. 10, October.

Orenstein J. and F. Manola[1988]: "PROBE Spatial Data Modeling and Query Processing in An Image Database Application." IEEE Transactions On Software Engineering. Vol. 14, No. 5, May.

Samet H.[1989]: " The Design and Analysis of Spatial Data Structures." Addison Wesley.

Sequoia 2000 Technical Report 93/11. Department of Electrical Engineering and Computer Science, University of California at Berkeley.

Stonebraker M. and M. Olson[1993]: "Large Object Support In POSTGRES." Sequoia 2000 Technical Report 93/30. Department of Electrical Engineering and Computer Science, University of California at Berkeley, May.

Wallace, G.[1991]: "The JPEG Still Picture Compression Standard." Communications of the ACM, Vol.34, No.4, April.

Wegener K, V. Lum, and C. Wu[1989]: "Image Management in a Multimedia Data Base System." Visual Database Systems, T. Kunii(Editor), Elsevier Science Publishers B. V. (North-Holland).

Wiederhold G. J. Brinkley, R. Samadani, and R. Clauer[1989]: "Model-Driven Image Analysis to Augment Databases." T. Kunii(Editor), Elsevier Science Publishers B. V. (North-Holland).

Zhou Q.[1989]: "A Methods for Integrating Remote Sensing and Geographic Information Systems" Photogrammetry and Remote Sensing Engineering. Vol. 55, No. 5, May.

Zhou W., W. Brunner and S. Wilkinson[1991]: "Full Integration of Remote Sensing Functionality into A Vector Based GIS." Proceedings of GIS/LIS'91, Atlantic, George, USA, Nov. 11.

Zhou W. and S. Wilkinson[1993]: "Large Heterogeneous Database Query and Visualization." Presented at GIS/LIS'93, Minneapolis, Minnesota, USA, Nov. 11.

method / size	JPEG (o.75)	RLC	12 bits RLC	RCL &Half Resolution
12MB (original)	1.2 MB	2MB	1.0MB	0.35MB

Table 1: A Scanned 24-bit Map File Sizes After Compression(in MB)

query / cost	one	two	three	four	five
old	38	240	15	162	423
new	3.5	2.8	2	2	27

Table 2: Various query Costs in Second

Traitement des objets spatiaux de grandes dimensions dans les systèmes d'information géographique

RÉSUMÉ

Traduction non disponible pour cause de livraison tardive du résumé définitif

LAND INFORMATION NETWORK FOR CANADA

DOUGLAS O'BRIEN, TERRY FISHER, BERT GUINDON, RICHARD BOUDREAU

Canada Centre for Remote Sensing, Surveys, Mapping and Remote Sensing Sector, 588 Booth Street, Ottawa, Ontario, Canada, K1A 0Y7

FRANÇOIS FAUCHER

Canada Centre for Mapping, Surveys, Mapping and Remote Sensing Sector, 615 Booth Street, Ottawa, Ontario, Canada, K1A 0E9

ABSTRACT

The Surveys, Mapping and Remote Sensing Sector (SMRSS) of Natural Resources Canada maintains over a dozen databases of geomatics information. Currently these databases are maintained and accessed separately. The Land Information Network for Canada (LINC) involves the development of a Geomatics Information Infrastructure (GII) and an Information Management Framework (IMF) to provide product and service integration for these databases.

The GII portion of LINC is the hardware and software that will fulfil client requirements for access to a fully integrated geomatics information system with built-in functionality for retrieving standard data products and generating custom products; increase effectiveness and efficiencies of geomatics information acquisition, management and dissemination, thereby reducing costs and duplication; enhance SMRSS's expertise in the field of geomatics data integration and dissemination and create a model for other organizations to consider in expanding this concept at a federal and national level; develop an industrial capability in geomatics infrastructure that can be exploited commercially in sales to domestic and foreign organizations requiring geomatics technology. The IMF provides a shell within which the GII works. This shell encompasses the policies, standards and planning required to implement LINC, as well as the business model and organizational structure needed to operate LINC.

A contract will be signed with private industry in mid 1994 for the development of a baseline LINC. This baseline system will be a limited implementation intended to be immediately useful by providing access to a partial on-line meta-data catalogue of SMRSS products, and allowing consistent access to a small number of on-line databases as well as off-line ordering and delivery of several existing products, and will be extendable so as to serve as a base to develop a full implementation. This baseline system will be operational in mid 1996.

KEY WORDS: Spatial Data, Information Infrastructure, Data Access

1. INTRODUCTION

In becoming more client driven, the Surveys, Mapping and Remote Sensing Sector (SMRSS) in 1991/1992 conducted a survey and cost-benefit study to define its essential topographic data requirements. Present and potential clients and other levels of government were surveyed. Based on the findings, a recommendation was formulated to build and maintain a comprehensive digital georeferenced database in which the topographic land base would be the most important component. Its structure and storage would be suitable for wide use by an extensive client base. An important aspect of this approach is to be the development of a fully digital system for handling all land related information across the Sector. This would allow clients to readily locate and access important land related information.

SMRSS accepted the recommendation and launched a project in November 1992 under the name of Total Digital Platform (subsequently renamed to the Land Information Network for

Canada - LINC) to develop a SMRSS Geomatics Information Infrastructure (GII) and Information Management Framework (IMF)

The objectives of LINC for SMRSS are to:

- fulfil client requirements for access to a fully integrated geomatics information system with built-in functionality for retrieving standard data products and generating custom products;
- increase effectiveness and efficiencies of geomatics information acquisition, management and dissemination, thereby reducing costs and duplication;
- enhance SMRSS's expertise in the field of geomatics data integration and dissemination and create a model for other organizations to consider in expanding this concept at a federal and national level, and;
- develop an industrial capability in geomatics infrastructure that can be exploited commercially in sales to domestic and foreign organizations requiring geomatics technology.

2. CURRENT DATA HOLDINGS

SMRSS currently maintains many geomatics data holdings. These data holdings are available externally as products, or are used internally to generate products for end users. Sales of SMRSS products are currently in the order of $7M per year and are expected to grow.

SMRSS data holdings are maintained separately by data custodians who have the mandate for their creation and update. These data will continue to be collected and maintained by the appropriate custodians. LINC will make all these geomatics data and their associated meta-data available to clients in an integrated view.

Geodetic Data

Geodetic positional control data are managed in an Oracle RDBMS environment as part of the National Geodetic Information System (NGIS) and are widely used in combination with imagery or vector topographic data for geometric correction. They are referenced to NAD83, use the geographic coordinate system and are encoded in ASCII. Privileged clients can already access these data holding via network (DATAPAC).

LINC will provide the tools required to access the geodetic data shown in Table 1.

National Topographic Data Base

Because of its wide use and general need for integration with other data types, National Topographic Data Base (NTDB) at 1:50,000 and 1:250,000 scale should be made available through LINC on its first release. NTDB data are well documented by an on-line catalogue and an off-line data dictionary, are georeferenced to NAD83 and use UTM as its coordinate system.

Satellite Imagery Meta-Data

Remote sensing imagery in raw form are stored digitally but are not available on-line because of the massive storage requirements. Radarsat International (RSI) has been contracted by SMRSS for distribution of processed or derived products. Table 2 shows the size and growth of the satellite image and meta-data archive.

Table 1: Geodetic Inventory Contents

Data Type	# Units
Primary Horizontal Control Networks	8,000 points
Primary Vertical Control Networks	85,000 points
Software Products	10
Geodetic Publication / Standards	50
Active Control System Data (ACS)	N/A
Global Position System (GPS) orbit data	N/A
Geoid Data	N/A

Table 2: CCRS Image Archive

Satellite/Sensor	# of scenes(Dec. 93)	# added/year
Landsat 1,2,3 MSS	85,000	0
Landsat 4,5 MSS & TM	380,000	30,000
SPOT 1,2,3 PLA & MLA	640,000	100,000
MOS 1,1B MESSR	35,000	6,000
NOAA AVHRR raw data	5,000	5,000
GEOCOMP AVHRR Products	1,000	1,000
ERS-1 SAR (passes)	3,000	1,000
Seasat SAR (passes)	200	0
Radarsat SAR (passes) (to be launched in 1995)	0	1,000

CCRS browse data are framed in the same manner as the inventory information, i.e. according to the standard framing system of the platform, WRS for Landsat, GRS for SPOT, etc.. Where there is no standard framing system, such as for NOAA AVHRR and ERS-1, the passes are arbitrarily broken into scenes. The meta-data for Seasat are provided in passes. For multispectral scanners three bands are used. Browse scenes are decimated to approximately 500 pixels by 400 lines. These decimated scenes are compressed using a standard JPEG lossy process. These compressed products are usually less than 50 kilobytes in size.

Canadian Geographical Names Data Base

The Canadian Geographic Names Database (CGNDB) is in digital form and stored on-line in a RDBMS (Oracle).

Work has continued to create a 'reduced' Internet accessible copy of the CGNDB. The database will reside on a completely independent server, an Internet node, which will be separated from the production CGNDB by a 'firewall' machine. The copy database, also an Oracle RDBMS, will be accessible by two distinct user groups: the first, members of the Canadian Permanent Committee

Table 3: Small Scale Map - Nameslists Available from Geographical Names Section

Scale	Map Title	Corresponding NAIS Map English	French	Bilingual	No. Rec.	No. Bytes
1:2 M	Atlantic Provinces	MCR 77	MCR 77F		2800	417647
	Quebec	MCR 42	MCR 42F		4000	596000
	Ontario	MCR 39	MCR 39F		3230	481270
	Prairie Provinces	MCR 27	MCR 27F		4404	656196
	British Columbia	MCR 3	MCR 3F		2732	407068
	Yukon and Northwest Territories					
	- East Half	MCR 134	MCR 134F		2361	351789
	- West Half	MCR 135	MCR 135F		2143	319307
1:7.5 M	Canada	MCR 128	MCR 128F		1673	249277
1:12.5 M	Canada	MCR 137	MCR 137F	MCR 138	714	106386
1:20 M	Canada	MCR 132	MCR 132F	MCR 139	314	46786
1:30 M	Canada	MCR 133	MCR 133F	MCR 136	128	19072

* A universal names list is provided for each map title at a given scale, regardless of the language version of the printed map.

on Geographical Names (CPCGN), will have access, using Oracle Forms, to all fields in their own region, but to core information fields only in all other regions; the second, will include libraries, the general public, and LINC, who will access the database via the World Wide Web. Approximately 350,000 records will be accessible to the public and LINC users for limited read only (free), or for customized purchase.

The CGNDB uses the 8-bit ASCII character set ISO 8859 which includes all French-accented characters but not all the diacritics now incorporated in some names in Canadian Native languages. Until ISO standards are adopted, these "hard-to-construct" characters will be represented numerically and surrounded by brace brackets.

{1}utselk'e	where {1} represents a barred L (i.e. an L with a forward slash (/) superimposed)
Dél{10}ne	where {10} represents an i with a reversed cedilla immediately below the letter

A number of ASCII text files are available which provide listings of those geographical names selected for several small scale National Atlas Maps (scale, title and MCR numbers shown in Table 3). These lists have been updated since publication of the maps, and will continue to be revised periodically.

Additional products listed in a catalogue may include Gazetteer of Canada Series, Geographical Names from the World Map, nameslists corresponding to other National Atlas Information Service maps of various scales, and lists of names for 1:250,000 and 1:50,000 scale National Topographic Series.

NTS Topographic Maps

Topographic maps (NTS) at 1:50,000 and 1:250,000 scales remain the best selling SMRSS products but are presently only available in hardcopy form making them impossible to access through LINC. However, as in the case of remote sensing imagery, the LINC meta-data should incorporate information from the topographic map lineage database and offer on-line ordering and off-line delivery.

Canada Lands Survey Information

Canada Lands Survey Records documents (CLSR) is a repository of official survey plans and related survey documents of Canada Lands. Meta-data for CLSR are maintained in the Survey Records Information System (SRIS), which is a RDBMS (Ingres) containing attribute information on CLRS documents. Data sets from the Registration, Survey and Land Management Triangular System (RESULT) are GIS/DBMS based information containing vector, attribute and some raster backdrop data defining the boundaries (extent) of cadastral parcels within Canada Lands.

Other Databases and Products

Other databases and products available from SMRSS that will be accessible via LINC include:

- Digital elevation data, which is maintained in a variety of formats such as Digital Terrain Elevation Data (DTED);
- National Atlas Information System (NAIS), which is maintained on a GIS and contains maps that comprise the National Atlas;
- National Air Photo Library (NAPL), which contains an air photo index and the physical air photos;
- Aerial Survey Database (ASDB), which contains photogrammetric control point information and other meta-data describing the contents of the air photo library;
- Canadian Aeronautical Charts (CANAC) including Canada Flight Supplement, VFR Charts, Canada Air Pilot, EnRoute and Terminal Area Charts;
- USA-Canada Boundary Description;
- National Cartographic Database (NCDB) and unstructured topographic data.
- Film Products.

3. IMPLEMENTATION

Phased Implementation

LINC will be implemented in phases. The ultimate goal, or Target Version, will be achieved through a progression of versions of LINC starting with the initial Baseline Version.

The Target Version of LINC incorporates all the elements required to meet client requirements. This includes: on-line access to all SMRSS digital geomatics data; seamless and feature based queries; visualization for most data types, and; a Canada wide client base.

Initially a Baseline Version will be developed in order to meet immediate client needs. This system will consider only a subset of the SMRSS resource holdings and required functionality in order to be timely.

Major Clients

Initial clients of the system will be clients internal to SMRSS who use other SMRSS data as part of their operations. Access to LINC will also be provided to internal digital distributors who would use the system to access, retrieve and forward the data through off-line methods (or other methods presently in use by those organizations) to their clients. Specifically, the Digital Distribution Office of the Products and Services Division would have this access.

In order to have the external clients' perspective and verify the system's performance over a WAN, the Baseline System will be accessible in certain organizations with existing spatial data infrastructures. This would allow those organizations to better use the baseline system and provide valuable feedback on its adequacy and potential problems that are network related.

4. LINC COMPONENTS

Information Management Framework

The Information Management Framework (IMF) is the set of policies and standards that provide a frame or context within which the Geomatics Information Infrastructure (GII) will operate. Together the IMF and GII form the Land Information Network for Canada (LINC), and both are equally important to the success of LINC.

Technology is not the only issue impacting on the effectiveness of information management within the Sector. The development and the adoption of standards also plays a crucial role. Useful and well integrated off-the-shelf and custom products will be accessed through the GII with confidence because of the use of a basic set of management principles. These principles, in turn, will help SMRSS meet its mandate in a more effective manner through the use of support policies and standards. This collection of policies and standards and the principles and mechanisms that will be required to develop and maintain them is referred to as the Information Management Framework. It is the responsibility of SMRSS to decide on and to act upon the IMF.

Policies and standards in the context of the IMF can be further broken down as follows:

Policies

- organization
- acquisition, use, dissemination, security and access control of data
- systems and applications
- pricing and intellectual property for products and services

Standards

- recognized standards for data exchange and modelling
- data communication, storage, description and administration
- geomatics data integration models (spatial and non-spatial)
- tool development and operation

Geomatics Information Infrastructure

The mission of the Geomatics Information Infrastructure system is to facilitate the production, exchange and integration of geomatics information so as to improve the internal efficiency of SMRSS and provide better external access to its products and services.

General requirements for the GII are:

- a meta-data catalogue describing the data sets and tool sets available via the GII, as well as data production and quality information;
- interfaces to a series of data holdings to allow the retrieval of data sets;
- user access facilities to allow users to identify, order and receive requested data sets;
- system management and accounting services;
- graphical services to order both digital

and non-digital data (e.g. paper or film products), as well as an upgrade path to accommodate data holdings that are migrated from non-digital to digital form;
- interface and translation services to export retrieved data sets to users' applications;
- support for appropriate data and meta-data standards;
- services and tools to support the Sector's various geomatics production processes;
- graphical services to display retrieved data;
- a system that can be extended by adding new services, further data sources, and additional users.
- data federations to allow unified views of data from disparate sources;
- support for third-party value-added services and tools;
- Canada-wide access to the system via Wide Area Networks;
- integrated access to data holdings outside of SMRSS; and
- tools for automatic maintenance and update of meta-data.

The GII consists of a number of components that interact via a communications backbone. The backbone is a set of networks (local, metropolitan and wide area) that together link the GII components. Users interact with the GII by means of client software. They refer to a meta-data catalogue that describes the data, tools and services available to them from the GII. Data and tool sets that are connected to the GII are called resource holdings. Data federations enable users to have integrated access to two or more related resource holdings. They do this by creating unified views that span resource holdings. Finally, there is a accounting system that handles all aspects of user account creation and management, as well as resource privileges and quotas, and data security and access control.

The GII works by, as much as possible, transforming data and services from connected resource holdings into a common environment called the common data model. In this way, data can be more easily integrated, and services that can operate in this common environment can be applied to any data within it. The common data model must support spatial data types and could be based on an existing spatial data exchange standard such as SAIF or SDTS. The important point is that it should be rich enough to embrace most existing spatial data formats.

Data translation is an important aspect of the GII. Translators are necessary in the servers of resource holdings to convert data and tools from the native data model of the holding to the common data model. Data translation is also necessary in the client software, where it must be possible for the GII to export retrieved data in format suitable for a user's application.

The GII catalogue contains the following information for navigating and using the GII:
- product and service guides describing the contents of resource holdings that are accessible from the GII, including their geographic extent, price, currency and quality information;
- production histories and schedules for products and services;
- examples of available products (especially raster data such as satellite radar and optical imagery) to give users an idea of what they will get should they order a particular product;
- index maps to allow users to identify for what areas they want products.

Each resource holding, either on or off-line, has entries in the catalogue that describe the data, the holding stores, as well as the tools that are available to act on the data.

User and Technical Groups

In order to ensure that LINC is properly developed a number of groups are overseeing it's development.

The LINC Steering Committee is a management level committee, consisting of SMRSS representatives, that is responsible for overseeing the LINC project. This includes all aspects of LINC in addition to the development of a Baseline System.

The LINC User Group consists of members from across SMRSS (and potentially external users) representing clients of LINC that are internal and external to SMRSS. The LINC User Group will provide input on LINC requirements, review LINC design including acting as beta testers, and report to the LINC Steering Committee.

The LINC Technical Group consists of members representing the resource holdings to be accessed by LINC. The LINC Technical Group provide the necessary technical information, as well as review all the technical proposals made throughout the project.

5. CONCLUSIONS

LINC will be the operational system to provide access to SMRSS geomatics products. It will be implemented in a series of phases, or versions, with the Baseline Version scheduled for operation in 1996. Initially, efforts will focus on providing access to meta-data on SMRSS geomatics products, but subsequent versions will provide increased access to the actual geomatics data holdings. In addition to the technology (i.e. hardware and software) required for such a system, the LINC project also includes the development of the necessary policies and standards that are required to make it a truly operational system. Efforts will be supported for the collection, maintenance and serving of the meta-data and geomatics data that will be available through LINC. A contract will be signed with private industry in mid 1994 for the development of a baseline LINC.

6. BIBLIOGRAPHY

Canadian General Standards Board, Working Group 4, "Proposed Standard Survey Form for Collecting Directory Information Describing Digital Geo-Referenced Datasets", 1993

Canadian Permanent Committee on Geographical Names, "Gazetteer of Canada" Brochure.

Committee on Earth Observation Satellites, Working Group on Data, Catalogue Subgroup, "Guidelines for an International Interoperable Catalogue System", Issue 2.1, April 1993.

Federal Geographic Data Committee, "Draft Content Standards for Spatial Metadata, November 3", November 3, 1992

Geographic Information Systems Division, SMRSS, "Delta-X: A Federated Spatial Information Management System".

Geographic Information Systems Division, SMRSS, "MetaView/GIS: A GIS Spatial Browser".

Surveys, Mapping and Remote Sensing Sector, "PlaNet - SMRS Products and Services Bulletin Board"

Total Digital Platform Project Team, Surveys, Mapping and Remote Sensing Sector, "Business Model and Requirements Analysis", May 1993.

Total Digital Platform Project Team, Surveys, Mapping and Remote Sensing Sector, "Target Vision and Migration Strategy", June 1993.

Le Canada sous réseau de télécommunication électronique

RÉSUMÉ

Le Secteur des levés, de la cartographie et de la télédétection (SLCT) de Ressources naturelles Canada assume la mise à jour de plus d'une douzaine de bases de données géomatiques. À l'heure actuelle, ces bases de données sont sollicitées et mises à jour individuellement. Le Canada sous réseau de télécommunication électronique (CARTE) nécessite la mise en place d'une infrastructure d'information géomatique et d'une structure de gestion de l'information de manière à obtenir des bases de données à intégration de produits et de services.

L'infrastructure du CARTE consiste en des matériels et des logiciels qui : répondront aux exigences des clients en matière d'accès à un système d'information géomatique entièrement intégré et dont les fonctions incorporées permettront d'extraire des produits d'information courants et de créer des produits sur mesure; rendront plus efficaces et plus rentables les opérations d'acquisition, de gestion et de diffusion de l'information géomatique, ce qui réduira les coûts et les risques de duplication; élargiront le champ des compétences techniques du SLCT dans le domaine de l'intégration et de la diffusion des données géomatiques et constitueront un modèle pour les autres organisations qui entendent développer ce concept à l'échelle fédérale et nationale; mettront en valeur le potentiel industriel de l'infrastructure géomatique dont l'exploitation à l'échelle commerciale se traduirait par la vente de la technologie géomatique aux organisations canadiennes et étrangères qui en ont besoin. L'infrastructure d'information géomatique s'articule autour d'une structure de gestion de l'information qui comprend les politiques, les normes et les activités de planification sur lesquelles repose la mise en oeuvre du CARTE, ainsi que le modèle commercial et la structure organisationnelle nécessaires au bon fonctionnement du CARTE.

On prévoit signer un contrat avec le secteur privé au début de 1994 pour la mise en place d'un réseau de base. En procédant à une mise en service partielle, on peut offrir sans délai un système fort utile; ainsi, les utilisateurs pourront accéder à un catalogue interactif partiel des métadonnées sur les produits du SLCT et accéder sur une base régulière à un petit nombre de bases de données en ligne ainsi qu'à des services en différé de commande et de livraison de plusieurs produits existants. Grâce aux possibilités d'enrichissement du système de base, on pourra bientôt envisager une mise en service complète. Ce système de base sera fonctionnel au milieu de l'année 1996.

SATELLITE DATA MANAGEMENT AND DISSEMINATION AT THE U.S. GEOLOGICAL SURVEY EROS DATA CENTER

Lyndon R. Oleson and Thomas M. Holm
U.S. Geological Survey
EROS Data Center
Sioux Falls, SD 57198

Darla J. Werner
Hughes STX Corporation
EROS Data Center
Sioux Falls, SD 57198

ABSTRACT

The U.S. Geological Survey (USGS) Earth Resources Observation Systems (EROS) Data Center has been collecting regional, continental, and global Earth observations data acquired by satellites, aircraft, and other information-gathering systems for more than 20 years. Currently, the Center holds more than 10 million satellite images and aerial photographs in both photographic and digital formats. In its role as the National Satellite Land Remote Sensing Data Archive, the Center is continually investigating and exploiting new technologies to ensure the long-term availability of these data. For example, the Center is converting its Landsat Multispectral Scanner (MSS) and Thematic Mapper (TM) satellite data from high-density instrumentation tapes to new, more stable archive media. In conjunction with the media conversion effort, the Center has successfully recovered more than 15,000 Landsat scenes from nearly 700 archive tapes affected by a condition called hydrolysis. The Center is also investigating new techniques to improve the accessibility of products and services from the archive. For example, the Landsat media conversion effort is providing the opportunity to collect digital browse images for user viewing through state-of-the-art information systems, such as the USGS Global Land Information System. Also, the Center is developing a capability for users to select, through an interactive graphical user interface, geographic or reduced-resolution subsets of very large digital satellite data sets. Such a capability will improve the effectiveness of their delivery by way of conventional wide-area networks.

KEY WORDS: Archive, Hydrolysis, Information System, Landsat, Media Conversion, Transcription

1. INTRODUCTION

The Earth Resources Observation Systems (EROS) Data Center is a data management, systems development, and research field center of the U.S. Geological Survey (USGS). The Center was established in the early 1970's to receive, process, and distribute data from National Aeronautics and Space Administration (NASA) Landsat satellites. The Center has been collecting regional, continental, and global Earth observations data acquired by satellites, aircraft, and other information-gathering systems for more than 20 years. The Center holds the world's largest collection of satellite- and aircraft-acquired images of the Earth's land surface. These holdings include more than 3 million satellite images and more than 7 million aerial photographs. The Center is also a major source for information about the holdings of foreign Landsat ground reception stations and data acquired by other countries' Earth-observing satellites.

The Center's management of global Earth observations data includes developing and operating advanced systems for receiving, processing, distributing, and applying land-related earth science, mapping, and other geographic data. More than 60,000 inquiries and orders are received annually, resulting in the distribution of more than 250,000 products to users around the world. Center activities include operating the National Satellite Land Remote Sensing Data Archive, a legislatively mandated responsibility, and the Federal Land Remote Sensing Research Program. The archive is responsible for maintaining a high quality data base of space acquired images of the Earth suitable for use in the study of global change and related scientific programs. The Federal Land Remote Sensing Research Program allows Federal agencies, universities, and other organizations to assign scientists and researchers to the Center on a full-time basis with complete access to its analytical equipment, data, and research facilities.

As a major participant in the U.S. Global Change Research Program, the Center provides data to scientists from around the world to improve

---

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

understanding and the ability to predict change. The Center is also participating in NASA's Earth Observing System (EOS) Program and plans to process and archive land-related data from sensors aboard the EOS polar-orbiting platforms. The Center was selected to serve as one of the principal EOS data archive and distribution facilities because of its demonstrated capabilities to archive, process, and distribute large volumes of data and related products from remote sensing satellite programs.

In its role as the National Satellite Land Remote Sensing Data Archive, the Center is continually investigating and exploiting new technologies to ensure the long-term availability of these data. For example, the Center is converting its current holdings of the U.S. Landsat archive to new, more stable archive media. This media conversion effort is providing the opportunity to assess the overall quality of the Landsat archive data and metadata and to collect digital browse images for user viewing through state-of-the-art information systems, such as the USGS Global Land Information System (GLIS). In addition, the Center is experimenting with image data subsetting and compression techniques to improve the overall effectiveness of the delivery of large-scale satellite images over conventional wide-area networks. Following are brief overviews of these examples of technology exploitation by the Center.

2. LANDSAT ARCHIVE MEDIA CONVERSION

Landsat data constitute the largest data set in the National Satellite Land Remote Sensing Data Archive. This data set includes more than 634,000 scenes of Landsat Multispectral Scanner (MSS) data, representing approximately 26 terabytes (TB) of data, that were acquired between 1972 and 1989 and are stored on 34,920 high density tapes (HDT). The archive also contains more than 214,000 Landsat Thematic Mapper (TM) scenes, representing approximately 41 TB of data, that have been acquired since 1982 and are stored on approximately 26,000 HDT.

To ensure the long-term availability of these data, the Center is converting these data to new, more stable archive media. In 1991, the Center funded the development of a specially configured system called the Thematic Mapper and Multispectral Scanner Archive Conversion System (TMACS) to transcribe the Landsat archive data to Digital Cassette Recording System Incremental (DCRSI) cassette tapes. The DCRSI tape drive implementation was selected because of its proven reliability, 48 gigabyte (GB) storage capacity, incremental buffering capability, and 107 megabyte (MB) per second transfer rate.

In addition to transcribing media, TMACS also supports data quality assessment and metadata generation and update operations. After the data have been transcribed to the DCRSI cassette, the tape is played back and a subset of the data is extracted. An orbital model is used to delineate individual scenes by calculating scene center latitude and longitude and the beginning and ending scan lines for each scene. Reduced resolution browse images are produced and visually assessed for cloud cover and image quality. The image data quality information is used to update archive metadata. The browse images are compressed and subsequently transferred to optical disks for access by users of the Global Land Information System.

Converting the HDT data to DCRSI cassettes improves the reliability of the data and saves archive space. When the conversion project is completed, more than 39,000 HDT's will have been transcribed to approximately 2,000 digital cassette tapes. The smaller physical dimensions of the DCRSI medium will allow the Center to achieve a 40-to-1 reduction in required archive storage space.

The Center has begun a program to monitor data integrity during the lifetime of the DCRSI tapes by using test patterns written at the beginning, middle, and end of each tape during data transcription. The DCRSI recorder's internal bit-error-rate (BER) test generator is used to write and then read the patterns and to compute BER statistics during verification processing. By establishing a baseline BER for each tape and maintaining a history of BER values through subsequent operations, the Center will establish quantitative quality criteria to help identify media degradation as early as possible. In addition, weekly system tests are performed on several prerecorded test tapes, allowing operations staff to verify, by comparing BER values, that the hardware is functioning at an acceptable level.

Conversion of the MSS data was begun in December 1992. Concurrent conversion of the TM data was begun in November 1993. At the current production rate, the MSS archive conversion will be completed in June 1994 and the TM data conversion should be completed by the end of 1996.

3. ARCHIVE DATA RECOVERY

A major problem encountered during the Landsat archive media conversion project was the recovery of data from tapes affected by hydrolysis. Hydrolysis is the process in which tape binders absorb atmospheric moisture that causes the binders to change chemically, creating a sticky residue buildup on the tape winder and cleaner tissues and rollers, as well as on the tape paths and heads of the tape recorders. This buildup can cause head wear, damage to the tape transport, physical damage to the tape, and the potential contamination of other tapes. Hydrolysis is exacerbated by humid conditions (while tapes are in storage or during transportation), but it also occurs at low humidity over extended periods of time, possibly as part of the natural aging process of older magnetic tapes. Most of the affected tapes at the Center were manufactured between 1978 and 1981 by two manufacturers. The hydrolysis process cannot be stopped; however, it can be temporarily reversed long enough to retrieve the data and copy it to new media.

Since April 1993, the Center has successfully recovered more than 15,000 Landsat scenes from nearly 700 affected tapes by applying low levels of heat (130° F) to the tapes for 24 hours to temporarily reverse the degrading condition. This technique, called baking, allows the tape to be played without physical damage to the tape or the drive. Local tests show that data can be recovered up to 11 weeks after a tape has been baked. In addition, it has been found that a tape can be baked several times and still achieve successful recovery.

4. THE GLOBAL LAND INFORMATION SYSTEM

The Global Land Information System (GLIS) is an online data directory, guide, and inventory system developed by the USGS to respond to the land data information and access needs of the U.S. global change research community. GLIS can best be characterized as a metadata system containing descriptive information about data sets and the associated functions that allow scientists to assess the potential utility of data sets, determine their availability, and place online requests for related data products.

The data set information in GLIS is maintained in three levels of detail -- the directory, guide, and inventory. The GLIS directory contains summary descriptions of entire data sets. Users can query the directory by using any combination of keywords (discipline, location, geophysical parameter), acquisition date, data source (spacecraft, sensor), geographic coverage, project, and investigator. The GLIS guides contain detailed descriptions of entire data sets, including such information as sensor specifications, extent of coverage, processing history, data quality, and product availability. The GLIS inventories contain detailed information about individual data set entities such as a Landsat scene. GLIS contains inventories of U.S. and foreign holdings of Landsat MSS and TM data, as well as the Center's holdings of National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) sensor data. GLIS supports temporal-, spatial-, and parameter-oriented queries of inventories, and the spatial query options include point, bounding rectangle, polygon, and geographic name searches.

In addition to providing the text-based directory, guide, and inventory query functions, GLIS also offers several graphical aids to users. Among these are capabilities for map-based geographic query specification and coverage plotting, and online digital image browse services for selected data sets. The map-based tools allow users to specify the area they wish to search by interactively "drawing" this rectangular or polygonal area on a map, or alternately by requesting to view the geographic coverage of their query results. The image browse services are primarily intended to allow users to visually judge overall image quality, determine the extent of cloud cover, and verify geographic coverage. GLIS currently contains more than 60,000 AVHRR browse images and soon will have the first set of Landsat MSS browse images being generated by the TMACS systems. These graphical aids are available to users through an X-windows-based, 32-bit graphics workstation interface called X-GLIS.

Although GLIS provides information about and access to a variety of regional, continental, and global land science data sets, these query and selection aids have proven particularly useful to those who are interested in selecting and ordering data from large satellite data inventories. For example, through GLIS a user can execute a query of the 600,000-scene Landsat MSS or the 60,000-pass AVHRR inventories, see the geographic coverage of individual scenes or passes plotted on a map, and request to view individual browse images for final selection and ordering.

5. GLIS USER ACCESS AND ASSISTANCE

Users who want to perform graphic-based queries can use an X-terminal or X-terminal emulator package on a PC.

From INTERNET: $TELNET xglis.cr.usgs.gov

Users who want to perform text-based queries can use an alphanumeric terminal or a terminal emulator package on a PC.

From INTERNET: $TELNET glis.cr.usgs.gov
or $TELNET 152.61.192.54

Direct Dial: Set modem to 8 bits, no parity, 1 stop bit Dial: (605) 594-6888

For information concerning system access and how to obtain a copy of the graphical interface software, contact GLIS User Assistance through one of the following:

Telephone: 1-800-252-GLIS (1-800-252-4547)
or commercial: (605) 594-6099

E-mail: GLIS@GLIS.CR.USGS.GOV

Mail: U.S. Geological Survey
EROS Data Center
GLIS User Assistance
Sioux Falls, SD 57198 USA

6. SATELLITE DATA DISTRIBUTION

Despite the dramatic improvements in digital data storage and transfer capacities that have been achieved over the past 20 years, data volume continues to be one of the most challenging aspects of satellite data distribution. Many of the same technological advances that have enhanced a data provider's ability to deliver large digital data products effectively and efficiently have also enhanced the producer's ability to generate such products. Factors such as greater spatial resolution and more spectral channels combined with powerful state-of-the-art computer systems and expanded communications band-widths, have all contributed to the generation of larger and larger satellite data sets.

One example of this challenge now facing the Center stems from its participation in NASA's EOS program. In preparation for its role as an EOS data archive, processing, and distribution facility, the Center is participating in early prelaunch activities to enhance the availability and utility of existing Earth science data. One such activity is the Global Land 1-km AVHRR Data Set project. This project is being conducted under the auspices of the NASA EOS Program in cooperation with the USGS, NOAA, the European Space Agency, and a worldwide network of AVHRR ground reception and data processing facilities. The primary objective of the project is to obtain complete 1-year coverage of all the land surfaces of the globe with 1-km AVHRR sensor data serving as a precursor data set for EOS investigators involved in the development of scientific algorithms and models.

One of the products being derived from the Global Land 1-km AVHRR Data Set is a 10-day global vegetation index composite that will allow scientists to experiment with vegetation index information on a global scale at the 1-km resolution. At full resolution, a single 10-day composite product represents 10 GB of data. The media being used to distribute these composites are 8-mm high-density cassette tapes. This medium was chosen because it required the fewest tapes and the least amount of operator involvement (see table 1). The time estimates provided in table 1 include computer write time only and do not include time for operator tasks such as tape mounting and dismounting and tape labeling. Therefore, even though the wall-clock time required to write a 10-GB composite to 8-mm high-density cassettes was 3 hours more than the time required to write the same composite to write-once compact disk (CD), it would be significantly more expensive operationally to handle and disseminate the 15 CD's. Also, it was felt that the users would prefer having a significantly smaller number of media.

Network delivery of a complete 10-GB composite is not a practical alternative. For example, assuming that a remote user has a moderately robust connection to the Internet (25-50 kilobytes per second), it is estimated that it could take from 2.3 to 4.6 days to transfer a full 10-GB composite. This scenario further assumes that the computer systems and network connections at both ends remain up throughout the period and that sufficient disk storage capacity is available to hold the data without operator intervention.

Assuming that most users would seldom require complete 10-day composites but would be interested in smaller geographic regions from one or more 10-day composites, the Center is exploring the use of data subsetting and compression techniques to provide a more effective network data delivery alternative. The Center is developing a capability for users to select, for network transfer, a geographic subset of a 10-day composite at a resolution of from 1-16 km. The selection would be done through a graphical user interface accessed by means of the EOS Data and Information System's interactive data query and order system, the Information Management System. A prototype of this capability should be available for demonstration and science-user feedback in July 1994, with a fully operational capability planned for October 1994.

Table 1: Media options for 10-GB composite product.

Media Type	Capacity	Number	Write Time
6250 bpi 9-track	.16 GB	63	16 hrs.
3480 cassette	.30 GB	34	9 hrs.
CD Write Once (6x)	.68 GB	15	3 hrs.
4mm cassette	2.0 GB	5	15 hrs.
8mm cassette (low)	2.3 GB	5	15 hrs.
8mm cassette (high)	5.0 GB	2	6 hrs.

7. SUMMARY

The EROS Data Center holds the world's largest collection of satellite- and aircraft-acquired images of the Earth's land surface, including more than 3 million satellite images and more than 7 million aerial photographs. In its role as the National Satellite Land Remote Sensing Data Archive, the Center is exploiting new technologies to ensure the long-term availability of these data. The Center is converting its Landsat MSS and TM satellite data from high-density instrumentation tapes to DCRSI cassette tapes. In conjunction with the media conversion effort, the Center has successfully recovered more than 15,000 Landsat scenes from nearly 700 archive tapes affected by hydrolysis. This recovery was accomplished by applying low levels of heat to the tapes to effect a temporary reversal of the hydrolysis.

The Center is also investigating and implementing new techniques to improve the quality and accessibility of products and services. For example, the Landsat media conversion effort is providing the opportunity to assess the overall quality of the Landsat archive data and metadata and to collect digital browse images for user viewing through state-of-the-art information systems, such as GLIS.

In preparation for its role as an EOS data archive, processing, and distribution facility, the Center is participating in early prelaunch activities to enhance the availability and utility of existing Earth science data. As part of these activities, the Center is exploring the use of data subsetting and compression techniques to provide an effective network data delivery capability for the 10-day global vegetation index composite products.

SELECTED REFERENCES

Cuddihy, E.F., 1980. The Aging of Magnetic Recording Tape. IEEE Transactions on Magnetics. Volume 16, Number 4, July 1980, p. 558.

Cuddihy, E.F., 1976. Hygroscopic Properties of Magnetic Recording Tape. IEEE Transactions on Magnetics. Volume 12, Number 2, March 1976, p. 126-135.

DeLancie, P., 1990. Sticky Shed Syndrome: Tips On Saving Your Damaged Master Tapes. In: Mix Magazine, May 1990, p. 148-152.

Holm, T.M., W.C. Draeger, R.R. Risty, 1993. Availability of Earth Observations Data From The

U.S. Geological Survey's EROS Data Center. In: Pecora 12 Symposium, Land Information from Space-Based Systems, Sioux Falls, South Dakota, August 1993, Proceedings: Falls Church, Virginia, American Society for Photogrammetry and Remote Sensing (in press).

Moulats, S., 1993. DCRSI Technology Overview. Ampex Corporation, Redwood City, California.

Oleson, L.R., 1992. The Global Land Information System. In: Earth and Space Science Information Systems, Pasadena, California, February, 1992, Proceedings: American Institute of Physics Conference Proceedings No. 283, p. 234-242.

Solving Sticky Tape Syndrome at EROS Data Center, 1993. In: NML Bits, Newsletter of the National Media Lab, National Media Lab, St. Paul, Minnesota. Vol. 3, Issue 2, September 1993, p. 6-7.

Werner, D., 1993. Status of the Landsat Thematic Mapper and Multispectral Scanner Archive Conversion System. In: Pecora 12 Symposium, Land Information from Space-Based Systems, Sioux Falls, South Dakota, August 1993, Proceedings: Falls Church, Virginia, American Society for Photogrammetry and Remote Sensing (in press).

Diffusion et gestion des données satellite
au Centre de Données EROS
de l'Agence Géologique Américaine

Lyndon R. Oleson et Thomas M. Holm
U.S. Geological Survey
EROS Data Center
Sioux Fass, SD 57198

Darla J Werner
Hughes STX Corporation
EROS Data Center
Sioux Falls, SD 57198

RÉSUMÉ

Le Centre de Données des Systèmes d'Observation des Ressources Terrestres (EROS) (CDE) de l'Agence Américaine de Géologie accumule vingt années de données d'observation de la Terre, de type régional, continental et global, obtenues à partir de satellites, d'avions et de tout autre système recueilleur d'informations. À l'heure actuelle, le CDE rassemble plus de 10 millions d'images satellites et aériennes sous les deux formats photographique et numéral. En tant que dépositaire des Archives Nationales de Données Terrestre Télédétectées et régisseur des données des sciences terrestres, le CDE cherche continuellement à développer et à exploiter de nouvelles technologies pour pouvoir assurer la disponibilité à long-terme de ces données. Par exemple, le CDE est actuellement en train de convertir les données provenant du Multispectral Scanner (MSS) et du Thematic Mapper (TM) des satellites Landsat, conservées sur des bandes à haute densité sur de nouveaux matériaux d'archivage. Ce projet de conversion nécessite la transcription de plus de 500,000 scènes MSS et TM, stockées sur environ 30,000 bandes numérales à haute densité, sur des cassettes magnétiques numérales à haute densité. Le CDE est aussi en train de développer et de mettre en place de nouvelles techniques pour améliorer la qualité et la disponibilité des produits et des services offerts par les archives. Ainsi, l'effort de conversion des données Landsat donne l'occasion d'évaluer l'état actuel des données dans les archives Landsat et de rassembler des images numérales affichées que les utilisateurs pourront accéder grâce à des systèmes d'information de pointe. L'effort de conversion s'accompagne de la mise en place d'un nouveau système de traitement des données qui améliorera les possibilités d'obtention de produits d'information standards à partir des archives. Le CDE est aussi en train d'expérimenter les possibilités de diffusion de l'imagerie satellite à grande échelle sur des réseaux à grande surface soit traditionnaux soit expérimentaux. Ces activités vont servir de précurseurs aux fonctions similaires que le CDE va devoir assurer en tant que l'un des principaux centre d'archives et de diffusion des données que le futur Système d'Observation de la Terre (EOS) va ressembler et auxquelles on pourra accéder au moyen du Système d'Information et de Données EOS.

OBTAINING EARTH OBSERVATION DATA FROM U.S. AND INTERNATIONAL DATA AND INFORMATION SYSTEMS

James R. Thieman

NASA / Goddard Space Flight Center, National Space Science Data Center,
Code 633.2, Greenbelt, MD 20771, USA.

Lola Olsen

NASA / Goddard Space Flight Center, Global Change Data Center,
Code 902, Greenbelt, MD 20771, USA.

ABSTRACT

Multiple archives of remote sensing data now exist, and the number is expected to increase in the years to come. It is becoming a more challenging task to determine which archives hold the data of interest, as well as to locate the relevant in-situ or correlative data and information. This accentuates the need for high-level directories and other information systems to aid in the search. Two organizations have been active in setting up systems to help users in locating data and information. First is the U.S. Interagency Working Group on Data Management for Global Change (IWGDMGC) composed of many federal agencies having data or information useful for the study of global change problems. This organization has created a computer-accessible Global Change Data and Information System (GCDIS) which is intended to aid users interested in the causes and effects of global change. The GCDIS includes many interconnected information systems which are useful for this purpose, such as the Earth Observing System Data and Information System (EOSDIS). Secondly, international information system activities are also ongoing through the auspices of the Committee on Earth Observation Satellites (CEOS). An International Directory Network (IDN) sponsored by CEOS has provided users in many nations with quick access to data information for the last several years. The IDN, in turn, leads to other network-accessible information systems which have detailed information about a wide variety of data. Examples of data information access within GCDIS, IDN, and related systems will be given.

KEY WORDS: Directory, Archives, Data location, Information systems

1. INTRODUCTION

The use of computer networks to locate and access data of interest is increasing rapidly. Online services are being offered by a variety of data archives and the trend is to make as many data products as possible directly accessible over the network. Remote locations throughout the world are now being effectively connected to the networks, and other locations that have network connections are depending more on it as their bandwidth increases, allowing larger files to be transferred efficiently. Some of the latest network search technologies, such as World Wide Web, Mosaic, WAIS, etc. can be characterized as experiencing explosive growth. The amount of information available through the network is overwhelming. It is often difficult to find information on a particular topic unless one knows exactly where to go. This does not help if one is searching for information on a particular subject but have no idea what location stores that information. Within the Earth observation/remote sensing community the locations of major data archives are well known, but there are a number of observing sites, archives, and data analysis locations around the world which may offer potentially valuable data as well. Several aids are now available which may help the user locate and use needed data. This paper will discuss the Global Change Data and Information System (GCDIS), the International Directory Network (IDN), and other tools which can be used to access Earth observation data.

2. THE GLOBAL CHANGE DATA AND INFORMATION SYSTEM

For many years now the U.S. federal agencies have been working together to support the data needs of the community in addressing the problems associated with global environmental change. The organization created to do this is called the Interagency Working Group on Data Management for Global Change (IWGDMGC). Most of the U.S. federal agencies which hold data relevant for the study of global change are a part of this group (Department of Agriculture, Department of Commerce, Department of Defense, Department of Energy, Department of Interior, Environmental Protection Agency, National Aeronautics and Space Administration, and the National Science Foundation. Together they have created the Global Change Data and Information System as a tool to aid the community in obtaining data and information relevant to global change. Detailed information about GCDIS can be found in the GCDIS

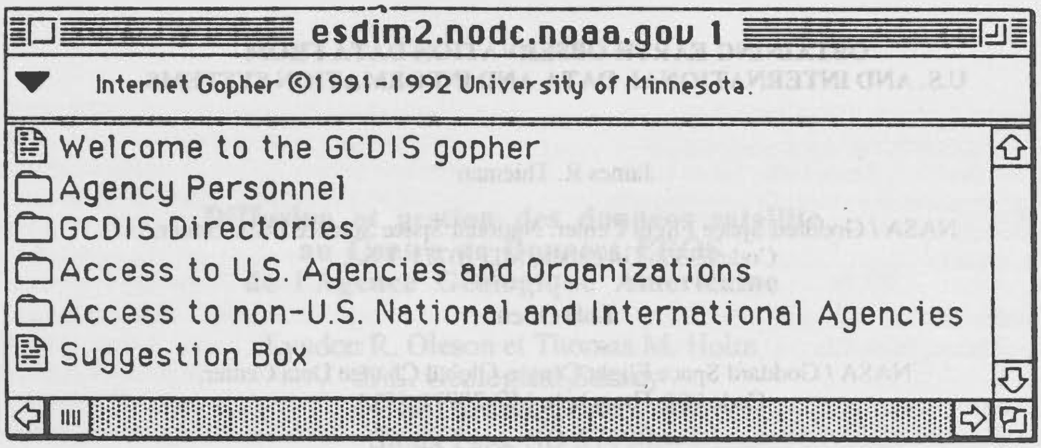


Figure 1: GCDIS Gopher Homepage

Program Plan (1992) and the Implementation Plan (1994).

The IWGDMGC member agencies received no additional funding to create a GCDIS. Consequently, the system had to make maximum use of existing services and capabilities. It was important that there be at least some initial GCDIS available as quickly as possible. Initiated on April 1, 1994, the system began with a wide variety of interconnected services united by the Gopher Internet search tool. It is intended that more sophisticated approaches be applied to GCDIS as time and resources permit, and as users have the opportunity to make their preferences known. Consequently, users are encouraged to access the system and give comments about their experience.

The GCDIS Gopher homepage (see Figure 1) is hosted by NOAA, and access to it is unrestricted. The GCDIS is a registered service available through the standard Gopher hierarchy starting at the University of Minnesota. One may also use the generic Gopher interface provided by NOAA through a direct logon to their system. To connect to this interface use the following sequence:

Telnet esdim2.nodc.noaa.gov
login: gopher

The GCDIS offers a large variety of services at no charge and many more are planned for the future. As mentioned previously, the GCDIS brings together a number of services which already existed within each agency. Many of the agencies already had Gopher servers operating. Where necessary, the servers available through the GCDIS have been tailored to highlight those services which are especially relevant to global change users. Examples of the types of services offered are:

- Detailed directories to data throughout NOAA
- A Geographic Information System section within the U.S. Geological Survey Gopher
- Information on climate data from the Department of Energy
- An extensive guide to EPA data and information called "Access EPA"
- A variety of socio-economic data and information useful for the interpretation of Earth observation data through the Consortium for International Earth Science Information Network (CIESIN)

Many other services could be listed, and others are being added. For example, in July the first public version of the NASA Earth Observing System Data and Information System will be made accessible through GCDIS. This includes specific information from the EOS Distributed Active Archive Centers (DAACs) dedicated to archiving and distributing data relative to observations of land, ocean, atmosphere, snow and ice, trace gases, etc.

Future plans also call for the use of other Internet tools for making GCDIS services readily accessible. World Wide Web servers already exist within many of the agencies. A Mosaic interface homepage for GCDIS would tie together the relevant services from each of the agencies, perhaps including access from pages oriented to specific topics.

This type of cross agency access to similar types of information available from multiple agencies is an important capability. One would like to be able to search all of these information sources simultaneously, but this is not easily implemented without someone taking the responsibility to make that possible. In the area of Earth observation data set descriptions, however, this work is already well underway through the International Directory Network.

3. THE INTERNATIONAL DIRECTORY NETWORK

Several locations in the GCDIS Gopher lead to directories such as the Global Change Master Directory, the NOAA directory, the CIESIN directory, and international directories in Italy, Japan, and Canada. All of these are part of the International Directory Network.

The International Directory Network (IDN) was created in an attempt to facilitate the sharing of information on Earth and space science data throughout the world . The IDN is composed of a federation of directory databases on widely-scattered computers which are interconnected through computer networks. Several of the directories are intended as a service to the entire world community and are made freely accessible to the community through computer network connections and dial-in lines. The user does not need to establish an account to use the directories and no passwords are required. Users guides to the directories are available (Shipe, 1993; Shipe and Bailey, 1993), but no training is necessary to use the system to get information on Earth and space science data sets as well as information on selected data information systems, spacecraft, sensors, and data gathering projects or campaigns.

Figure 2 is a picture of the current and potential new nodes of the IDN. Two types of nodes are shown. The coordinating nodes have completely identical databases at each location (U.S., Italy, and Japan). They are open to use by the general community and are particularly intended for access by users in the continental vicinity in which they are located (i.e., America, Europe, and Asia). Access information for these nodes is supplied in Table 1 for a variety of mechanisms. The other cooperating nodes share information with the coordinating nodes but have databases which may be a subset of those of the coordinating nodes. In some cases, such as in Canada, the cooperating node maintains a complete copy of the database. Cooperating nodes contribute information to the overall network but may only be used by a small group for local data management and information purposes. Many of these types of nodes exist now and more are being considered for the future as indicated by the dashed lines to some of these nodes. Access information for both coordinating and cooperating nodes is summarized in the brochure entitled "International Directory Network". These brochures are available from the authors.

The glue which binds the IDN is a standard method of describing dataset information called the Directory Interchange Format (DIF). All dataset information exchanged among the nodes is transferred in this form. Contributions of information on datasets by the general community are encouraged. Creation of a DIF file describing a dataset is a relatively simple process and, once this file is created, it can be automatically loaded into the directory databases and shared with the other directory nodes. Thus, the information is quickly spread throughout the world. The Directory Interchange Format Manual (1993) is the definitive document for this format.

The IDN provides more than just dataset information, however. When the user accesses a particular dataset description there is sometimes an indication that an automatic connection is available to another information system which has more detailed information about this dataset. Whereas directories just have brief, overview information about datasets, the other information systems may offer more complete information about the dataset such as calibration information, sensor characteristics, detailed usage information, etc. These types of information systems are classified as "guide" systems. Still others contain details about the elements of the

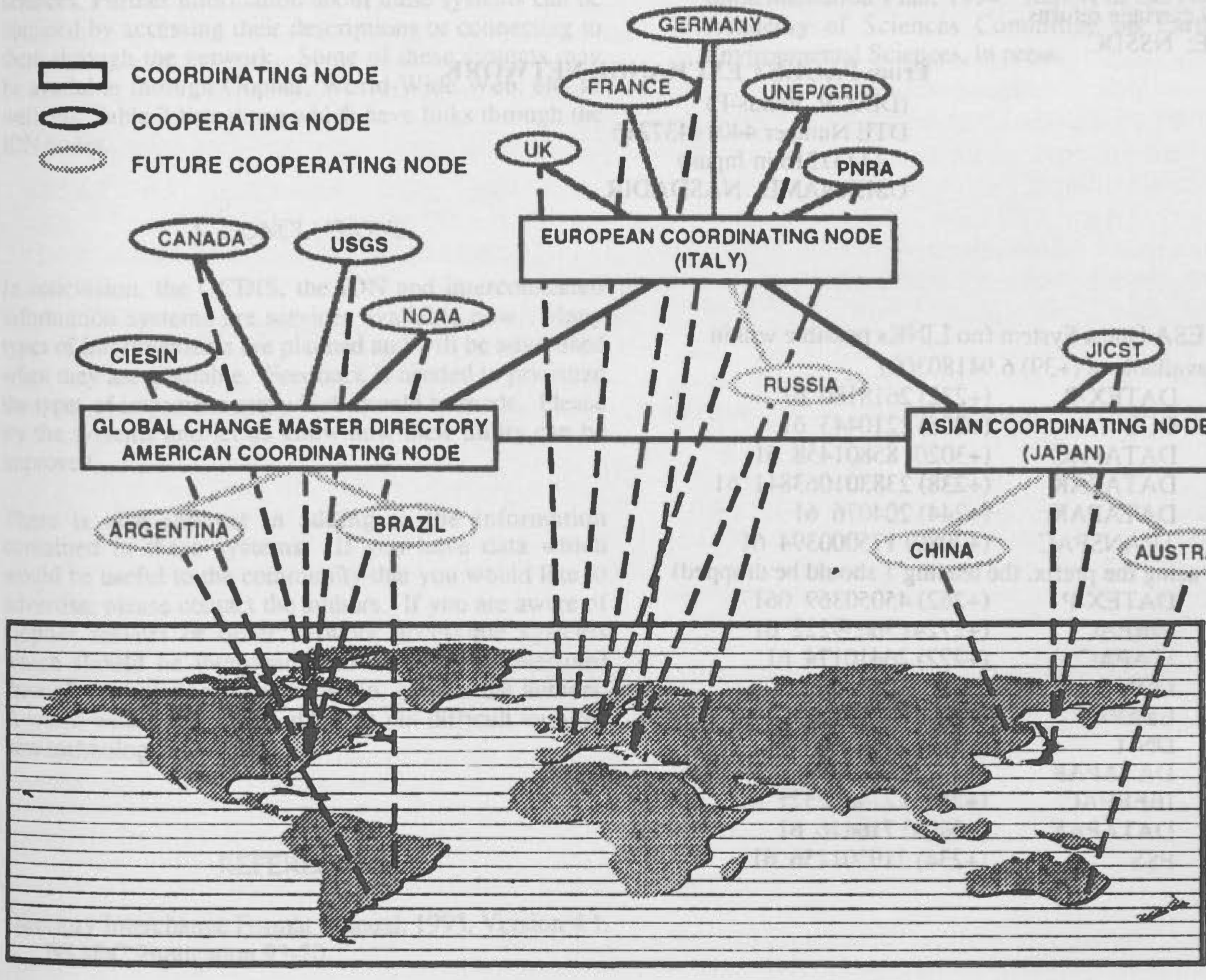


Figure 2: The International Directory Network

TABLE 1. DIRECTORY ACCESS

To use the coordinating nodes of the IDN one can follow the procedures listed below according to what type of access (network connection or dial-in line) is available.

American IDN Node Access	Japanese IDN Node Access	European IDN Node Access
	From INTERNET	
Telnet GCMD.GSFC.NASA.GOV USERNAME: GCDIR Telnet NSSDCA.GSFC.NASA.GOV USERNAME: NSSDC	Telnet 133.56.72.1 USERNAME: NASDADIR	Telnet 192.106.252.160 USERNAME: ESAPID
	From NSI/DECnet (SPAN)	
$ SET HOST NSSDCA USERNAME: NSSDC	$ SET HOST 41950 USERNAME: NASDADIR	$ SET HOST 29628 USERNAME: ESAPID
	By DIAL-IN LINES	
Set to 8 bits, no parity, 1 stop bit Dial 301-286-9000 CONNECT 1200 (or 2400 or 300) Enter several carriage returns ENTER NUMBER MD CALLING 55201 (or 55202) CALL COMPLETE Enter several carriage returns USERNAME: NSSDC	Set to 8 bits, no parity, 1 stop bit XON/OFF Dial 81-492-94-6400 (0492-96-6400 in Japan) 1200–9600 bps, (DEC Kanji) USERNAME: NASDADIR	Set to 8 bits, no parity, 1 stop bit Dial (+39) 6 9417335 * (See alternative telephone list below)
	From PACKET EXCHANGE NETWORK	
	(DDX-P, Venus-P) DTE Number 44014437216 (4437216 in Japan) USERNAME: NASDADIR	

* File 21 of ESA Quest System (no LINKs possible within Quest; help available at (+39) 6 94180300)

Austria	DATEX-P	(+232) 2618180 61
Belgium	DCS	(+206) 2210443 61
Canada	DATAPAC	(+3020) 85801458 61
Denmark	DATAPAK	(+238) 238301063841 61
Finland	DATAPAK	(+244) 204076 61
France	TRANSPAC	(+2080) 175000394 61
(note: when using the prefix, the leading 1 should be dropped)		
Germany	DATEX-P	(+262) 45050369 061
Ireland	EIRPAC	(+2724) 36059222 61
Italy	ITAPAC	(+222) 26410174 61
	DDN	111005306009
	ESAPAC	299020030001
NL	DN-1	(+204) 1290176 61
Norway	DATAPAK	(+2422) 110627 61
Spain	IBERPAC	(+2145) 214062321 61
Sweden	DATAPAK	(+2403) 710416 61
UK	PSS	(+234) 219201156 61

Table 2 Earth-observation-relevant information systems connected to the IDN

Acronym	Title
DALI	CNES-SPOT IMAGE Catalogue
EOC	Earth Observation Center, Japan
ESA EARTH IMAGES	ESA Earthnet online catalogue (formerly LEDA)
GISS	Goddard Institute for Space Studies
IRPS	Image Retrieval and Processing System
NCDC	National Climatic Data Center
NSF	NSF Science and Technology Information System
OCEANIC	Ocean Network Information Center
RESTEC	Remote Sensing Technology Center of Japan
SDCS	SAR Data Catalog System
SMRSS	Surveys, Mapping and Remote Sensing Sector
UA - GEODATA CENTER	University of Alaska Fairbanks/GeoData Center
URI AVHRR ARCHIVE	University of Rhode Island AVHRR Archive

datasets such as whether there are data available for a particular location and/or time period. Systems which have dataset "granule" information are classified as "inventories". Some information systems have combinations of directory, guide, and/or inventory capabilities. They may also allow the user to browse and manipulate the data.

Table 2 is a list of some of the Earth-observation-relevant data systems which have connections through the directory and may be helpful to a researcher in the Earth sciences. Further information about these systems can be obtained by accessing their descriptions or connecting to them through the network. Some of these systems may be available through Gopher, World-Wide Web, etc. as well, but Table 2 lists those which have links through the IDN nodes.

4. CONCLUSIONS

In conclusion, the GCDIS, the IDN and interconnected information systems are services available now. Many types of improvements are planned and will be advertised when they are available. Feedback is needed to prioritize the types of improvements which should be made. Please try the systems and let us know how their utility can be improved.

There is also interest in adding to the information contained in these systems. If you have data which would be useful to the community that you would like to advertise, please contact the authors. If you are aware of Gopher servers or other network accessible systems which should be interconnected or at least advertised through the existing systems, again, contact the authors. A world-wide cooperative effort is not difficult with the new technology capabilities.

REFERENCES

Directory Interchange Format Manual, 1993. Version 4.1, NSSDC Publication 93-20.

Shipe, Janis L., 1993. Users' Guide to the Master Directory-2 Alphanumeric Interface, NSSDC.

Shipe, Janis L. and P. A. Bailey, 1993. Quick Reference Guide to the Master Directory-2 Alphanumeric Interface, NSSDC.

The U.S. Global Change Data and Information System Program Plan, 1992. Report of the National Academy of Sciences Committee on Earth and Environmental Sciences.

The U.S. Global Change Data and Information System Implementation Plan, 1994. Report of the National Academy of Sciences Committee on Earth and Environmental Sciences, in press.

Comment obtenir des données d'observation de la Terre à partir de systèmes américains et internationaux de données et d'information

Résumé

Il existe maintenant d'abondantes archives de données de télédétection, lesquelles devraient continuer de s'accumuler dans les années à venir. Une telle manne d'informations oblige cependant les chercheurs à déterminer et à repérer les données dont ils ont besoin, ainsi que tous les renseignements complémentaires portant sur le site d'observation considéré. Une telle tâche ne saurait être menée à bien sans l'aide de répertoires de haut niveau ou d'autres systèmes d'information. Deux organisations se sont jusqu'à présent employées à mettre au point des systèmes pour aider les utilisateurs à repérer les données et l'information nécessaires. Mentionnons d'abord la U.S. Interagency Working Group on Data Management for Global Change (IWGDMGC), qui regroupe plusieurs organismes fédéraux disposant de données ou d'informations utiles à l'étude de problèmes de changement global. Cette organisation a créé le Global Change Data and Information System (GCDIS), exploitable par ordinateur, qui vise à aider les utilisateurs qui s'intéressent aux causes et aux effets des phénomènes de changement global. Le GCDIS comprend lui-même plusieurs systèmes d'information interconnectés et spécialement conçus à cette fin, comme par exemple le Earth Observing System Data and Information System (EOSDIS). Par ailleurs, les données du système d'information international sont également accessibles par l'entremise du Committee on Earth Observation Satellites (CEOS). Ce comité a à son tour créé le International Directory Network (IDN), qui fournit aux utilisateurs de nombreux pays un accès rapide aux données accumulées depuis quelques années. Ce système IDN oriente ensuite les chercheurs vers d'autres systèmes d'information en réseau, leur donnant ainsi accès à des quantités d'information dans des domaines variés. Le présent document fournit également des exemples de moyens d'accès à l'information, par les systèmes GCDIS et IDN et d'autres systèmes connexes.

INTEGRATING DIFFERENTIAL GPS WITH AN INERTIAL NAVIGATION SYSTEM (INS) AND CCD CAMERAS FOR A MOBILE GIS DATA COLLECTION SYSTEM

El-Sheimy N. and K. P. Schwarz

Department of Geomatics Engineering, The University of Calgary
2500 University Drive, N. W., Calgary, Alberta, Canada T2N 1N4
ISPRS Commission II, WG II/1

KEY WORDS: Highway survey system, Sensor integration, Geographical information system (GIS) , CCD cameras, Inertial navigation system , Differential Positioning with GPS.

ABSTRACT

The creation of a Geographic Information System (GIS) for road networks requires large amounts of data which currently are obtained by manually digitizing existing maps with some update from conventional photogrammetric and surveying techniques. In order to collect digital data faster, the use and integration of digital sensors is required which work in kinematic mode, i.e. from a moving vehicle. The University of Calgary has developed a precise mobile survey system for road inventory and general GIS applications. The system integrates a cluster of CCD cameras, a GPS receiver, and an Inertial Navigation System (INS), to automatically collect data in a road corridor at velocities of 50 - 60 km/h and to store this data in a GIS system format. The updated GPS/INS information is used to geometrically correct the images collected by the CCD cameras which record all details along the highway within a corridor of about 50 m. The shutters of the cameras and the output of the INS system are synchronized by the clock of the GPS receiver.
In this paper an overview of the sensor integration and the mathematical transformations required to convert the image coordinates into 3-D coordinates will be given. Also, some data storage, merging and manipulation problems will be addressed. Special emphasis is given to the contribution of each subsystem to the overall error budget of the derived 3-D coordinates which have been assessed using test results recently obtained.

1. INTRODUCTION

During the last decade, the demand for GIS in management and design applications has greatly increased. The realization of a GIS in many application areas still suffers from data acquisition problems. To be of value to the user, a GIS must be updated regularly, so that the information in the data base correctly represents the real world. However, the acquisition of up-to-date GIS data by conventional survey techniques is prohibitive in cost and has therefore limited the applicability and usefulness of GIS to potential users. Described in this paper is an attempt to overcome this problem.

A data acquisition system has been designed and implemented at The University of Calgary (U of C) that can be used to selectively update GIS data bases very quickly and inexpensively. The system integrates a cluster of Charged-Coupled-Devices (CCD) cameras, an Inertial Navigation System (INS), and satellite receivers of the Global Positioning System (GPS). Figure 1 shows the system in schematic form. The overall objective was the development of a precise mobile highway survey system that could be operated at a speed of 60 km per hour and achieve an accuracy of 0.3 m (RMS) with respect to the given control and a relative accuracy of 0.1 m (RMS) for points within a 50 m radius. This accuracy is required in all environments including inner cities, where a stand-alone GPS is not reliable. Sensor integration has been optimized to reach the requirements of the survey market. The data flow has been streamlined to facilitate the subsequent feature extraction process and transfer into a GIS system. The new system is named **VISAT** and derives its name from the fact that it utilizes Video images, an INS system, and the GPS Satellite system. For further details on the design of the VISAT system, see Schwarz et al (1993a) and El-Sheimy and Schwarz (1993); for the workstation design, see Li et al (1994).

This paper gives an overview of the most important components of the VISAT system, discusses the underlying principle of extracting three dimensional (3-D) coordinates from the video images through the use of INS/GPS data, highlights some of the integration problems, and analyses results recently obtained in field and laboratory tests.

2. SYSTEM CONCEPT AND FUNCTIONALITY

The hardware component of the VISAT system consists of three sensor systems, two ASTECH PXII GPS receivers, a LTN 90-100 strapdown INS, and two COHU 4912 CCD cameras.

The GPS is capable of providing very accurate position and velocity under ideal conditions. However, such conditions do not often exist. Independent GPS navigation requires at least four satellites. To reach the accuracy required for the VISAT system, it has to be used in differential mode, e.g. using double differencing techniques. The major drawback of

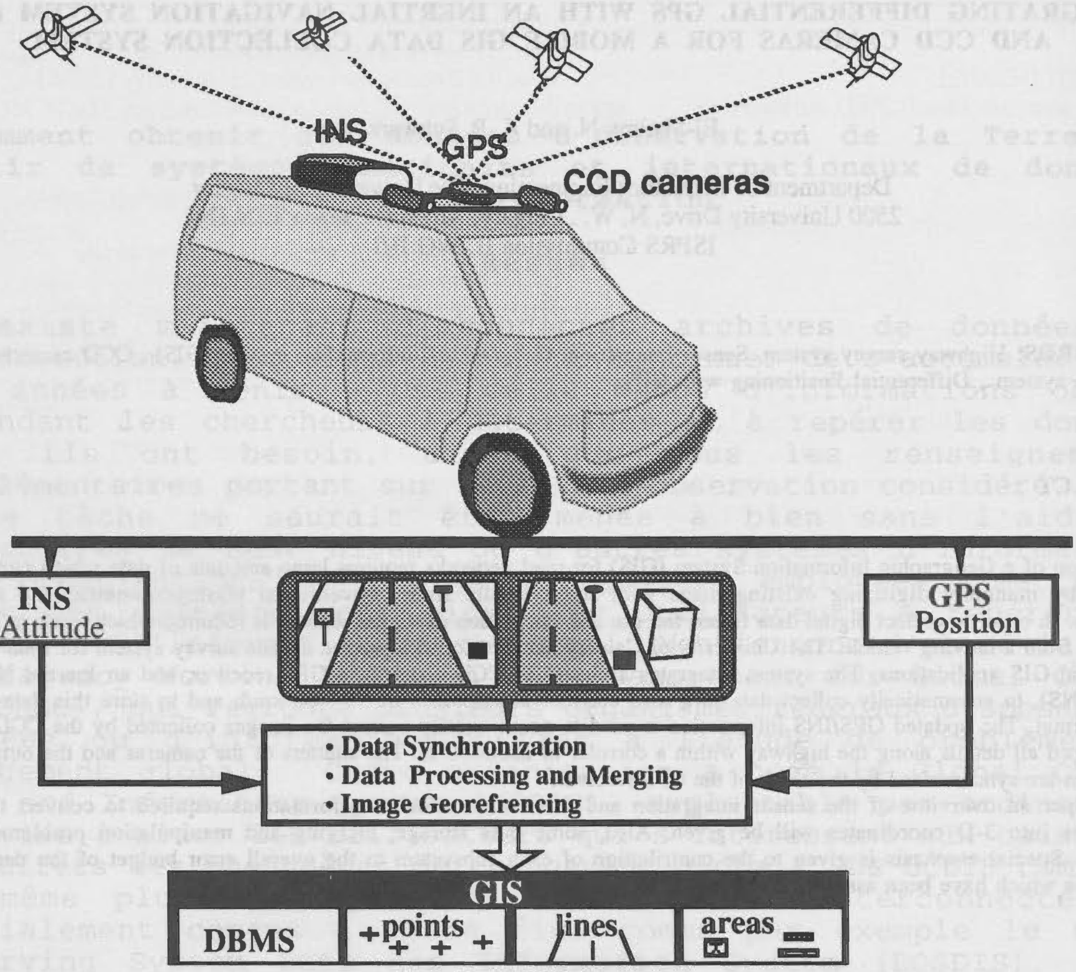


Figure 1: The VISAT system

GPS is the accuracy degradation due to poor satellite geometry, cycle slips, satellite outages, and dynamic lag during maneuvers.

The INS measures linear acceleration and angular rates very accurately and with minimum time delay. For short time intervals, the integration of acceleration and angular rate results in extremely accurate velocity, position, and attitude with almost no noise or time lags. However, because the INS outputs are obtained by integration, they drift at low frequencies. To obtain very accurate outputs at all frequencies, the INS should be updated periodically using external measurements. One of the typical external measurements is the Zero-Velocity-Update (ZUPT) which is simply obtained by stopping the vehicle. Disadvantages of using ZUPTs are that:

• the system will be limited to semi-kinematic applications;

• on highways and high traffic roads, it is not possible to stop the vehicle without interrupting the traffic flow;

• the production rate, which could be critical in many projects, will be reduced.

The integration of GPS and INS, therefore, provides a navigation system that has a superior performance in comparison with a stand-alone system. For instance, GPS derived positions have approximately white noise characteristics over the whole frequency range. The GPS derived positions and velocities are therefore excellent external measurements for updating the INS, thus improving its long term accuracy. Similarly, the INS can provide precise position and velocity data for GPS signal acquisition and reacquisition after outages. This reduces the time and the search domain required for detecting and correcting cycle slips. To optimally combine the GPS and the INS data, a Kalman filtering scheme is used (Schwarz et. al., 1990). The University of Calgary has developed a decentralized Kalman filter software KINGSPAD (KINematic Geodetic System for Positions and Attitude Determination) for processing INS/GPS data. The GPS data are Kalman filtered to obtain estimates of position and velocity which are then used as quasi-observations for the INS Kalman filter. At the same time, the GPS data are continuously checked for cycle slips. For more details on the mathematical formulation and Kalman filtering alternatives, see Wei and Schwarz (1990a and 1990b). KINGSPAD can perform the following functions:

• processing the data in three different modes, that is, pure GPS, pure INS, and hybrid INS/GPS.

• defining which GPS data will be used to update the INS, namely, position, velocity, and position/velocity,

• viewing individual space vehicle (SV) data, thus allowing the rejection of specific SV in the GPS processor,

• selecting the GPS update rate according to a specific application (airborne, land application),

computing the updated INS position, velocity, and attitude at 1-64 Hz to suit different applications

• applying rapid static integer ambiguity resolution techniques for short baselines (under 7 km).

• applying 'on-the-fly' ambiguity techniques for kinematic processing of GPS data.

The output from the KINGSPAD program will be used to georeference (position and orient) the CCD images as described in section 4.

CCD cameras produce standard video signals (e.g. RS-170, CCIR). The CCD camera clusters used in the VISAT system provide three dimensional positioning with respect to the VISAT reference which in most cases is the perspective center of one of the cameras. Recent trends in CCD cameras are characterized by increased resolution and improved radiometric quality. Use is made of sensor chips of resolution 4k by 4k, with standard chips ranging from 400 to 580 lines and 500 to 780 sensor elements.

Determining the VISAT reference in three-dimensional space is in principle a problem of trajectory determination. This is equivalent to finding the six parameters that define rigid body motion in space, i.e. three translations and three rotations. The GPS provides three positions and three velocities, while the INS provides three positions (velocities), and three rotations. Therefore, the problem of determining the VISAT reference can be seen as determining six parameters out of 12 independent measurements.

3. HARDWARE CONFIGURATION

In the vehicle, the three sensors are interfaced to a regular PC-AT, which controls the different tasks through programmed interrupt processes. The PC has been designed and assembled at the UofC with assistance by Logical Solution Ltd. in Calgary. Figure (2) shows a block diagram of the VISAT system hardware interfaced to the PC. The VISAT hardware is controlled by a real-time data logging software VIG (Video, Inertial, GPS) designed at the UofC. The VIG has been written in such a way that in can manipulate different hardware components. The output files from the VIG software are the input to KINGSPAD software.

The VIG software performs the following functions through an interrupt process:

• storing the INS position and attitude information at a data rate of 64 Hz. The INS is interfaced to the computer through the ARINC-429 serial board. At present the board issues 64 interrupts per second through IRQ 5;

• storing the GPS position and velocity at a data rate of 2 Hz (2 k byte/sec). The GPS receiver is interfaced to the computer through a smart serial board. With the smart board the GPS data can be grabbed once per second, thereby reducing the number of GPS data interrupts;

• storing the satellite ephemeries for post processing of the GPS data;

• interrupting the computer time chip through the IRQ 7 upon receiving the PPS pulse from the GPS receiver. The PPS pulse can be activated through any parallel port depending on user requirements;

• storing any special events, e.g. start/end of ZUPT, switch off/on the cameras etc., using the keyboard interrupts through IRQ 1. All the interrupts throughout IRQ 1 are stored on a file with time tag to allow post processing;

• grabbing and storing two or three simultaneous images (0.5 M byte) every 0.4 sec. The program allows different recording intervals for the cameras. The two CCD cameras are interfaced with the computer through the MATROX IMAGE-CLD frame grabber and the MATROX IMAGE-640 baseboard. The IMAGE-640 works as a temporary storage and an interface between the frame grabber and the host. The IMAGE-640 has an EISA bus interface which provides a full 32-bit data

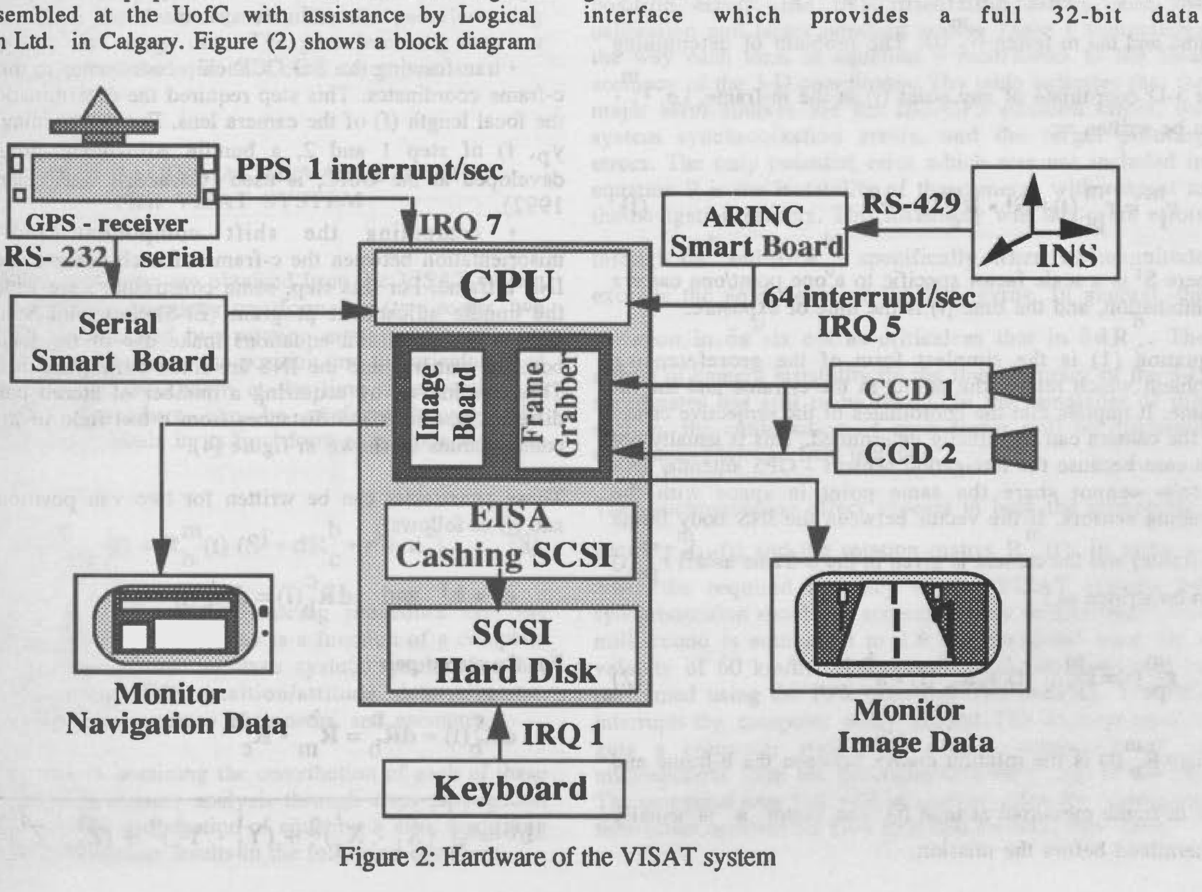


Figure 2: Hardware of the VISAT system

interface to the host and therefore allows fast image transfer. The images are stored in such a way that all images grabbed at the same time will have a header file that contains the necessary information for georeferencing them. This header will be filled after processing the INS and GPS data using the KINGSPAD software;

• guiding the operator by a user friendly monitor which contains all the important information such as video images, position, velocity , attitude, number of grabbed images, the computer disk space, the active hard disk, number of locked satellites, satellites that have dropped or are affected by cycle slips etc. ;

• switching the cameras off and on in order to limit the storage requirements. This feature is very important in situations when the vehicle has to stop, for example at traffic signals.

4. GEOREFERENCING OF VIDEO IMAGES

The problem of georeferencing video images can be defined as the problem of transforming the 3-D coordinate vector r^c of the camera frame (c-frame) to the 3-D coordinate vector r^m of the mapping frame (m-frame) in which the results are required. The m-frame can be a system of curvilinear geodetic coordinates (latitude, longitude, height), a system of UTM or 3TM coordinates, or any other earth-fixed coordinate system. For more details on georeferencing of remotely sensed images using INS/GPS data see Schwarz et al (1993b).

The process of georeferencing the video images includes the determination of the camera projective centers (p.c.) in the m-frame, i.e. $r^m_{pc}(t)$, and the rotation matrix between the c-frame and the m-frame $R^m_c(t)$. The problem of determining the 3-D coordinate of any point (i) in the m-frame, i.e. r^m_i, can be written as:

$$r^m_i = r^m_{pc}(t) + S^i \cdot R^m_c(t) \cdot r^c \qquad (1)$$

where S^i is a scale factor specific to a one point/one camera combination, and the time (t) is the time of exposure.

Equation (1) is the simplest form of the georeferencing problem which relates the points in the c-frame and the m-frame. It implies that the coordinates of the projective center of the camera can be directly determined. This is usually not the case because the navigation sensors - GPS antenna/ INS gyro's- cannot share the same point in space with the imaging sensors. If the vector between the INS body frame (b-frame) and the camera is given in the b-frame as a^b, $r^m_{pc}(t)$ can be written as :

$$r^m_{pc}(t) = r^m_{INS}(t) + R^m_b(t) \cdot a^b \qquad (2)$$

where $R^m_b(t)$ is the rotation matrix between the b-frame and the m-frame measured at time (t). The vector a^b is usually determined before the mission.

Similarly, the INS b-frame (gyro frame) cannot be aligned with the c-frame. If the small constant misorientation dR^b_c between the two frames is obtained, $R^m_c(t)$ can be written as:

$$R^m_c(t) = R^m_b(t) \cdot dR^b_c . \qquad (3)$$

The computation of 3-D coordinates of any feature includes the following major steps :

• transforming the 2-D image coordinates measured in the computer coordinate frame to the 2-D image coordinates of the camera CCD chip. This step requires the determination of the camera principal point coordinates (x_p, y_p). Figure (3) shows the relation between the two systems.

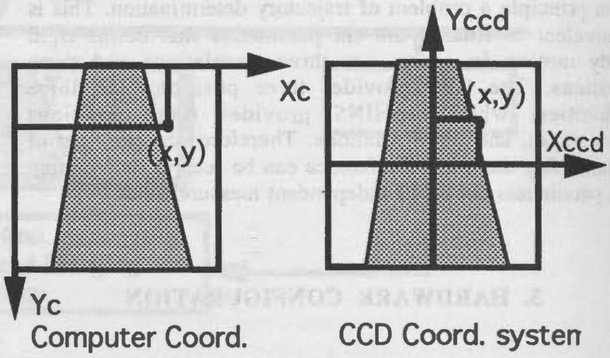


Figure 3 The relation between the computer and CCD chip coordinate systems

• transforming the 2-D CCD chip coordinates to the 3-D c-frame coordinates. This step required the determination of the focal length (f) of the camera lens. For determining (x_p, y_p, f) of step 1 and 2, a bundle adjustment program, developed at the UofC, is used (Cosandier and Chapman 1992).

• computing the shift component and the misorientation between the c-frame of each camera and the INS b-frame. For this step, some constraints , are added to the bundle adjustment program (El-Sheimy and Schwarz 1993). The constraint equations make use of the fact that both the cameras and the INS are fixed during the mission. This is achieved by acquiring a number of stereo pairs at different positions and distances from a test field of ground control points as shown in figure (4).

These constraints can be written for two van positions (i) and (j) as follows:

$$b^i = b^j \quad \text{and} \quad dR^c_b(i) = dR^c_b(j) \qquad (5)$$

For the stereo-pair (i),

$$dR^c_b(i) = dR^c_b = R^b_m \cdot R^m_c \qquad (6)$$

$$b^{(i)} = \sqrt{(X^i_l - X^i_r)^2 + (Y^i_l - Y^i_r)^2 + (Z^i_l - Z^i_r)^2} \qquad (7)$$

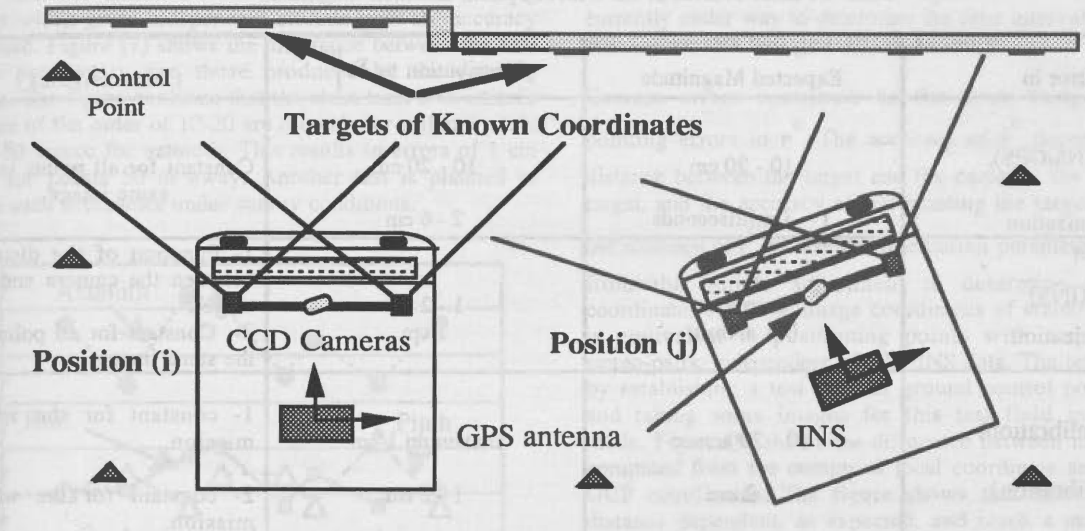


Figure 4: The test field for calibrating the VISAT system

Where (X_l^i, Y_l^i, Z_l^i) and (X_r^i, Y_r^i, Z_r^i) are the left and the right camera perspective coordinates of image(i) as obtained from the bundle adjustment, and R_c^m is the rotation matrix between the c-frame and the m-frame as obtained from the bundle adjustment.

- interpolating the perspective center coordinate of the cameras and the attitude information at the time of exposure. This requires the camera to be synchronized with the GPS/INS data.
- merging the data necessary for the georeferencing process with the image data. The georeferencing data are contained in the header in front of each stereo-pair

5. ACCURACIES ACHIEVABLE WITH THE VISAT SYSTEM

The object coordinates are obtained from the VISAT system through a fourteen parameter transformation (two scales, two translation vectors, and two rotation matrices) from the c-frame to the m-frame. As the position and the orientation of the two cameras are known at the time of exposure, an intersection procedure can be performed using the georeferencing formula in its final form as given in equation 8

$$r_i^m = r_{INS}^m(t) + R_b^m(t)(S^i \cdot dR_c^b \cdot r^c + a^b). \qquad (8)$$

It is clear from the georeferencing procedure that the accuracy of the 3-D coordinates is a function of a complete processing chain. This, involves system synchronization, GPS positioning, INS position/attitude determination, system calibration, accuracy of cameras, and geometry.

As a first step in obtaining the contribution of each of these factors, a standard error analysis through error propagation is performed. The propagation of equation 8 after neglecting the second order terms results in the following equation:

$$\delta r_i^m = \delta r_{INS}^m(t) + \delta R_b^m(t) \cdot (S^i \cdot dR_c^b \cdot r^c + a^b) + R_b^m(t) \cdot (\delta S^i \cdot dR_c^b \cdot r^c + S^i \cdot \delta dR_c^b \cdot r^c + S^i \cdot dR_c^b \cdot \delta r^c + \delta a^b) \qquad (9)$$

Equation 9 contains three major group of errors that contribute to the final accuracy of the 3-D coordinates derived from the VISAT system. They are the INS/GPS position errors, the INS orientation errors, and the calibration and target pointing errors. Table 1 summarizes the way each term in equation 9 contributes to the final accuracy of the 3-D coordinates. The table indicates that the major error sources are the INS/GPS position errors, the system synchronization errors, and the target pointing errors. The only potential error which was not included in equation 9 is the instability of the cameras with respect to the navigation sensors. This instability will add more errors through δa^b and δdR_c^b, specifically when its magnitude exceeds the noise level of the INS gyros. In general, the variation in δa^b is not as critical as that in δdR_c^b. The effects of camera instability on the final accuracy of the 3-D coordinates has still to be tested. In the remainder of this section, the contribution of each factor will be discussed using recent lab and field test data.

The synchronization effect appears in both the interpolated vector $r_{INS}^m(t)$ and the rotation matrix $R_c^m(t)$. In order to reach the required accuracy of the VISAT system, the synchronization should be accurate to few milliseconds (one millisecond is equivalent to 1.6 cm positional error for a velocity of 60 km/h). The system synchronization will be performed using the PPS from the GPS receiver. The PPS interrupts the computer every second. The interrupt handler gets a computer time tick with a resolution of 53 microseconds from the programmable time chip of the PC. The computer time tick will be used to solve the ambiguous time offset between the GPS time and the computer time.

Table 1 : Error sources of the VISAT system and their magnitude

Error in	Expected Magnitude	Contribution to δr_i^m	Characteristics
• First term :			
1– δr^m_{INS} (INS/GPS).	10 - 20 cm	10 - 20 cm	Constant for all points in the same image.
2- Synchronization	1 - 3 milliseconds	2 - 6 cm	
• Second term			1- Function of the distance between the camera and the target.
1– δR^m_b(t) (INS)	1 arcmin	1 - 2 cm	
2- Synchronization	0.5 arcmin	1 cm	2- Constant for all points in the same image.
• Third term			
1- dR^b_m (Calibration)	10 - 20 arcsec	maximum 1 cm	1- constant for the whole mission.
2- δa^b (Calibration)	1-2 cm	1 - 2 cm	2- constant for the whole mission.
3- δr^c(Target pointing)	One pixel	1 - 8 cm	3- Function of the pixel size.

Therefore, all time tags can be put in the GPS time frame every time a call is made to the PC timer. One problem with this technique, is the synchronization of the CCD cameras. The problem arises from the fact that most of the CCD cameras do not issue an interrupt indicating the time of exposure. Therefore, it should be synchronized mathematically using the specifications of the Video rate. The maximum error in synchronizing the cameras is of the order of 3.0 milliseconds as shown in figure (5). Results in this figure were calculated using the following equation:

Synchronization error = t2 - (t1 + 1/30)

where:

t1 is the computer time before grabbing the images;
t2 is the computer time after grabbing the images;
1/30 is the time of grabbing one image.

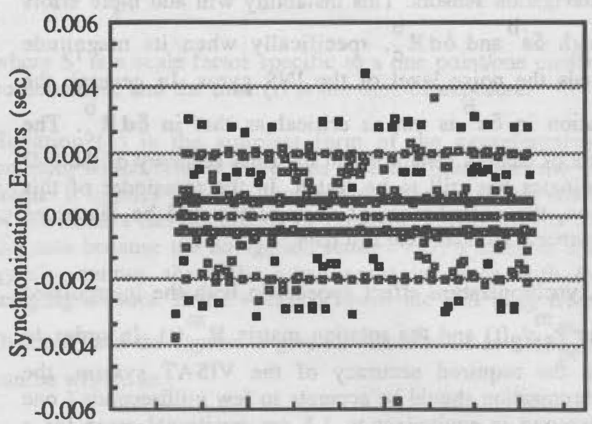


Figure 5 : Camera Synchronization Errors.

The GPS contributes to the georeferencing process by being the main positioning component for the perspective center vectors r^m_{pc}(t) of the two cameras. For precise position determination, differential GPS is needed. A variety of differential kinematic truck surveys using GPS have been performed over the last few years. Most of these tests have shown some limitation due to low data rate, outages due to satellite masking which cause poor geometry, and loss of carrier phase lock. Therefore, the use of GPS position and velocity will be limited to updating the INS data stream. The motivation for the integration of INS and GPS is to exploit the benefits of each system. The achievable accuracy of INS/GPS for many road tests is at the level of a few centimeters when a consistent satellite constellation is available. Typical INS/GPS positional errors for a truck survey over a baseline of 16 km is shown in figure (6). Agreement with pre-surveyed control points was generally better than 10 cm with an RMS of 7 cm.

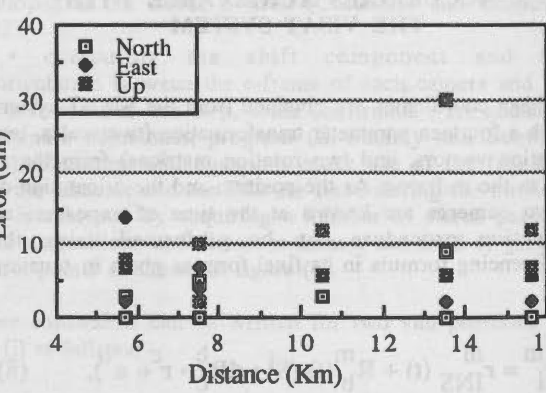


Figure 6: GPS/INS Positional Errors.

The INS contributes to the georeferencing process by determining the position and attitude of the c-frame, i.e. by giving the vector r^m_{pc}(t) and the attitude matrix R^b_m(t). For attitude determination, the INS is the primary system. In order to reach the required accuracy from the VISAT system, the derived INS roll, pitch, and azimuth should be at the level of one arcminute. To check such accuracies, lab tests were conducted at the U of C Mechanical Engineering Department.

In this test the Litton 90-100 was tested on a rotating platform which gives independent attitude with an accuracy of 3 arcsec. Figure (7) shows the difference between the INS attitude parameters and those produced by the rotating platform. The figure confirms that the short term INS attitude errors are of the order of 10-20 arc seconds for roll and pitch, and 30-50 arcsec for azimuth. This results in errors of 1 cm or less for points 50 m away. Another test is planned to confirm such accuracies under survey conditions.

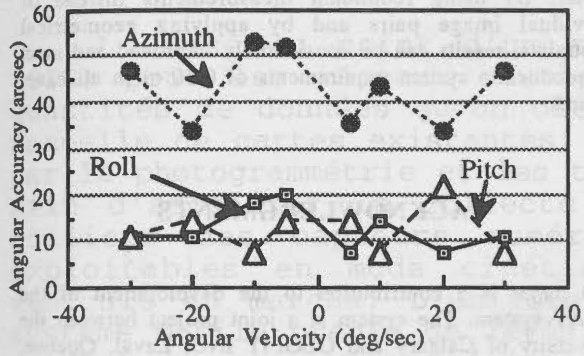


Figure 7: INS Attitude Accuracy as a Function of Angular Velocity

The updated INS positions will be mainly used for interpolating camera coordinates and detecting GPS cycle slips. Inertial systems combine very high short-term accuracy and high data rate, and, thus, are ideally suited for these tasks. The accuracy of using INS positions in interpolating the camera coordinates depends on the position drift rate between updates. Figure (8) shows a typical drift behavior between update intervals of 20 sec in static mode. The figure confirms a drift of 3 cm/10 sec in latitude. This means that for a GPS updating interval of 1 second, the INS position accuracy is at the level of few millimeters, better than the GPS accuracy for this time interval.

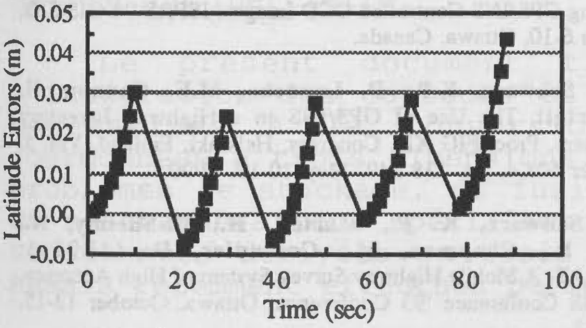


Figure 8 : INS position drift in static mode

The idea of detecting cycle slips is based on comparing the measured integer cycle number with the number predicted from the INS derived position. If a cycle slip occurred, the respective carrier phase can be corrected by resetting the integer value to the integer value closest to the predicted value. Obviously, the method will work best for short time intervals. In environments were cycle slips are frequent, as for instance inner cities, the success of cycle slip detection depends mainly on the INS position accuracy. In order to correct cycle slips at the one cycle level, INS positions must be accurate to 0.5 cycle, or better, for the outage interval. The accuracy of the positions derived from INS are a function of the update rate and the vehicle dynamics. Detailed tests are currently under way to determine the time interval in which a stand-alone INS will give 10 cm accuracy.

Camera errors contribute to the error budget through pointing errors in r_c. The accuracy of r_c depends on the distance between the target and the cameras, the size of the target, and the accuracy of pin-pointing the target. To check the accuracy of r_c we use the orientation parameters obtained from the bundle adjustment to determine the object coordinates from the image coordinates of stereo pairs. This is equivalent to positioning points with respect to the stereo-pairs, independent of GPS/INS data. The test was done by establishing a test field of ground control points (GCP) and taking some images for this test field in kinematic mode. Figure (9) shows the difference between the distances computed from the computed local coordinate and from the GCP coordinates. The figure shows that these errors are distance dependent, as expected, and reach a magnitude of 10-15 cm for objects 50 m away from the van.

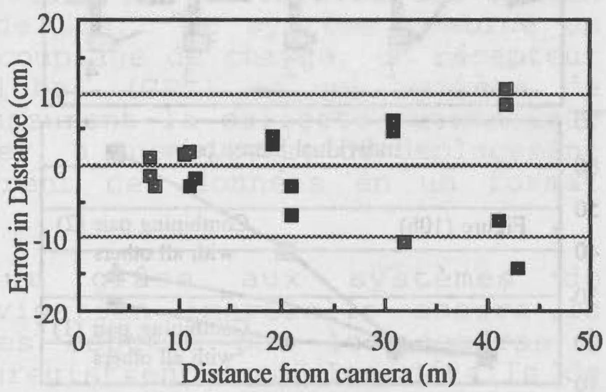


Figure 9 : The accuracy of measuring distances with CCD cameras

The georeferencing equation (5) contains four unknowns (3 coordinates and one scale factor) in three equations. Having a second camera of the same scene will add three extra equations and one unknown for the same point (i). A least squares solution of the space intersection between the two rays from the two images is computed. With a one second image data rate and a speed of 60 km/h, the van moves 16 m between exposures, and it is expected that the same object will , therefore, appear in four consecutive images. This adds extra redundancy and geometry to the spatial intersection problem. Figure (10) illustrates the error reduction resulting from adding redundancy and geometrical constraints to the minimum configuration. Figure (10a) shows the geometry of the situation with the target point visible in four pairs of images. If the four individual pairs are used for positioning, the upper curve in Figure (10b) is obtained. It shows errors which grow rapidly with distance from the target and are considerably larger than expected. This is almost certainly due to stability problems of the camera mount in this specific run. When combining all images to determine the target point, the positioning error is reduced to 0.25 m which is well within the required accuracy. It should be noted that the combination is not just a weighted mean of the four individual determinations but indirectly introduces the coordinate differences between exposure stations, measured by GPS/INS, as a constraint into the equation. Instead of combining all images, selected images can also be used. The two lower curves in Figure (10b) show this situation for two pairs of images. In general this will be sufficient to achieve the desired accuracy.

As mentioned above, the large errors in the results from individual pairs of images are most likely due to an instability of the camera mount, i.e. in a $\mathbf{dR_b^c}$- matrix that changes with time. This can be either caused by differential vibrations between camera and INS or by physical movement of the complete camera mount. The latter is most likely the case. In this specific test, a different vehicle was used and it was detected later on that the mounting was much less stable than usually.

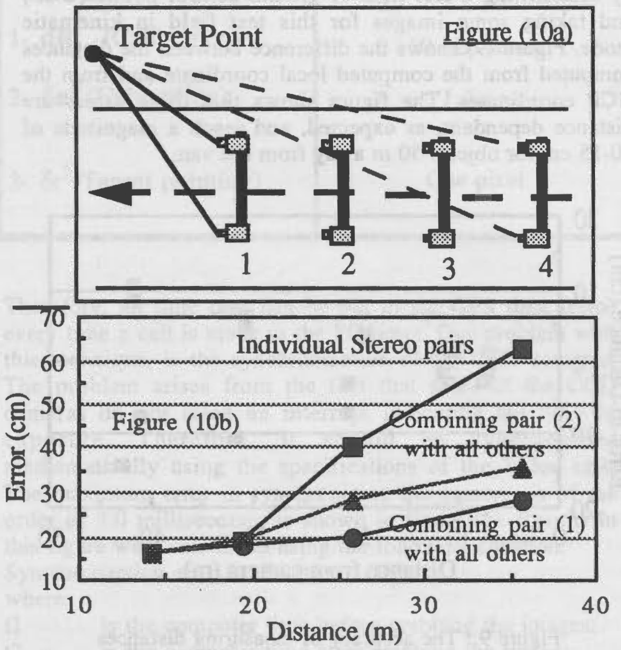


Figure 10 : Effect of Geometrical Constraints on 3-D coordinates

6. CONCLUSIONS

The prototype of the VISAT system presented in this paper is the first mobile GIS data acquisition system of this accuracy class. It is designed to work with comparable accuracy in rural as well as in urban areas where GPS may not be available all the time. To maintain a consistent accuracy of 0.30 m (RMS), the integration and mutual control of INS and GPS is an essential aspect of the system concept. The complementary features of INS and GPS permit the resolution of cycle slips and outages, as well as INS position and attitude control. The tight integration and synchronization of the GPS/INS component with the CCD camera cluster results in the precise georeferencing of each image, which is a necessity for a system of this accuracy class and gives considerable flexibility in data processing.

A major focus of the paper is the analysis of recent field tests, designed to estimate the contribution of each individual error source to the total error budget. Georeferencing can be done with a position accuracy of 5-10 cm (RMS) for the individual exposure station and an attitude accuracy of 10-20 arc seconds for pitch and roll and 30-50 arc seconds for azimuth. The position results were obtained from road tests, the attitude results from lab tests for angular velocities of up to 30 degrees per second. The imaging component, which provides positioning with respect to the current van position, achieves an accuracy of 10-15 cm for distances of up to 50 m. Further improvements can be expected with a third camera being added to the camera cluster. By using redundant measurements instead of individual image pairs and by applying geometrical constraints, results can be considerably improved and meet the production system requirements of 0.30 m in all cases studied.

ACKNOWLEDGMENTS

This paper is a contribution to the development of the VISAT system. The system is a joint project between the University of Calgary and GEOFIT INC, Laval, Quebec. Darren Cosandier is gratefully acknowledged for his support during system testing.

REFERENCES

[1] **Cosandier, D., Chapman, M. A.**, High Precision Target Location for Industrial Metrology, Videometrics, SPIE OE/Technology, Boston, November, 1992.

[2] **El-Sheimy, N., Schwarz K.P.**, Kinematic Positioning In Three Dimension Using CCD Technology, VNIS93 Conference, Ottawa, October 12-15 1993.

[3] **Li, R., M. A. Chapman, Qian, L., Xin. Y., and K. P. Schwarz**, Rapid GIS Database Generation Using GPS/INS Controlled CCD Images, ISPRS 94 GIS/SIG, June 6-10, Ottawa, Canada.

[4] **Schwarz, K.P., D. Lapucha, M.E. Cannon, H. Martell**, The Use of GPS/INS in a Highway Inventory System. Proc. FIG XIX Congress, Helsinki, Finland, Vol. 5, paper 508.4, pp. 238-249, May 10-19, 1990.

[5] **Schwarz, K. P., Martell, H., El-Sheimy, N., Li, R., Chapman, M., Cosandier, D.** (1993a): VISAT- A Mobile Highway Survey System of High Accuracy, VNIS Conference '93 Conference, Ottawa, October 12-15, 1993.

[6] **Schwarz, K.P., Chapman, M.A, Cannon, M. W., Gong, P.** (1993b): An Integrated INS/GPS Approach to the Georeferencing of Remotely Sensed Data, PE&RS Vol. 59, No. 11, November 1993, pp. 1667-1674..

[7] **Wei, M. and K.P. Schwarz** (1990a): Testing a Decentralized Filter for GPS/INS Integration. Proc. IEEE PLANS 1990, Las Vegas, pp. 429-435, March 20-23, 1990,

[8] **Wei, M. and K.P. Schwarz** (1990b): A Strapdown Inertial Algorithm Using an Earth-Fixed Cartesian Frame. Navigation, Vol. 37, No. 2, pp.153-167, 1990.

Intégration des appareils de prise de vues à DCC, du GPS en mode différentiel et d'un INS en un système mobile de collecte de données de SIG

RÉSUMÉ

La mise au point d'un système d'information géographique (SIG) pour les réseaux routiers nécessite l'apport de grandes quantités de données qu'on obtient actuellement par numérisation manuelle de cartes existantes, corrigées jusqu'à un certain degré par la photogrammétrie et les techniques de levés traditionnelles. Afin d'accélérer la collecte des données numériques, il faut utiliser des capteurs numériques à fonctionnement intégré, exploitables en mode cinétique, c'est-à-dire à partir d'un véhicule en marche. L'université de Calgary a mis au point un système mobile précis de levés, destiné à l'inventaire des routes ou à des applications générales de SIG. Ce système combine un groupe de caméras à dispositif de couplage de charge, un récepteur de système de positionnement global (GPS) et un système de navigation inertielle (INS), qui assurent la collecte automatique des données dans le corridor routier, à une vitesse de déplacement de 50 à 60 km/h et qui enregistrent ces données en un format exploitable par SIG.

L'information mise à jour grâce aux systèmes de positionnement global et de navigation inertielle assure la correction géométrique des images captées par les caméras à couplage de charge, lesquelles enregistrent tous les détails de chaque côté de la route, dans un corridor d'environ 50 mètres. L'horloge du récepteur GPS synchronise le déclenchement des obturateurs des caméras avec les signaux de sortie du système de navigation inertielle.

Le présent document traite tout particulièrement de l'intégration des capteurs et des transformations mathématiques nécessaires à la conversion des coordonnées d'images en coordonnées à trois dimensions. On y abordera également certains problèmes de stockage, de fusion et de manipulation des données. On s'attachera tout spécialement enfin à la précision globale des coordonnées dérivées en trois dimensions, que les résultats des essais récemment effectués permettent d'évaluer.

DESIGN AND SIMULATION OF A REAL-TIME MAPPING SATELLITE FOR THE KINGDOM OF SAUDI ARABIA

Abdulaziz I. Alobaida and Kurt Novak

Military Institute of Surveying and Geographical Studies
P.O. Box 89088, Riyadh 11682
Saudi Arabia
Tel.: ( ) ; Fax: ( )

ABSTRACT

The goal of this interdisciplinary research is to design a mapping and remote sensing satellite system (to be named SULTAN) for the Kingdom of Saudi Arabia (KSA) and its neighbouring regions and to prove its mapping feasibility by a simulation study. It will be equipped with state-of-the-art remote sensing, positioning, and attitude determination technology that is either currently available or will be available in the near future. The satellite system should provide the users with spatial data capable of producing maps and map substitutes of scale 1/25,000 without the need for ground control.

Two options are suggested and analyzed -- a system of one satellite or of two satellites. In the case of a one-satellite system, a threefold stereoscopic line scanner sensor will be investigated. The second mapping system consists of two satellites, both equipped in the above similar manner. The orbits of both satellites are similarly inclined and separated by about 2.6 degrees in the right ascension of their ascending nodes. These orbits were specifically chosen such that an area on the equator of the Earth would be covered by the vertically looking sensors of the first satellite and about 30 degrees, off the nadir, titled sensors on the second satellite. With this configuration, a specific scene can be covered by six images, three from each satellite.

The orbits of the system will be as low as possible, sun-synchronous, possibly minimum-drift orbits, nearly circular and polar. The positions of the satellites will be determined by on-board GPS receivers, while the sensors' attitude will be determined by one or more stellar sensors.

The general form of the collinearity equation is used in which a linear and higher order functions interpolate between orientation images to relate image coordinates, sensor parameters, and ground coordinates.

The point positioning accuracy is investigated by a variance-covariance analysis to ensure that the required accuracy is achieved. In the case of the one-satellite system, using image measurement accuracies of one-tenth of a pixel, GPS positioning accurate to 5 cm, and stellar sensor measurements good to 2 arc seconds, yield 1.2, 2.0, and 3.9 meters in the X, Y and Z directions respectively. As a result of this study a configuration of the real-time satellite system will be proposed that yields best accuracies and high reliability without the need for ground control points.

MISE AU POINT ET SIMULATION D'UN SATELLITE DE CARTOGRAPHIE EN TEMPS RÉEL POUR L'ARABIE SAOUDITE

RÉSUMÉ

Cette recherche interdisciplinaire a un double but : élaborer un système satellitaire de cartographie et de télédétection (qui sera dénommé SULTAN) destiné à l'Arabie saoudite et aux régions voisines et vérifier ses capacités cartographiques par une étude de simulation. Ce système bénéficiera des dernières innovations en matière de télédétection, de positionnement et de détermination de l'attitude, déjà disponibles ou en passe de le devenir. Il devra fournir aux utilisateurs les données spatiales nécessaires à l'établissement de cartes ou de substituts de cartes à une échelle de 1/25 000, sans contrôle sur le terrain.

Les deux solutions proposées et analysées sont un système à un satellite et un système à deux satellites. Dans le cas du système à un satellite, l'étude portera sur un détecteur à balayage linéaire stéréoscopique triple. Le deuxième système comprend deux satellites dotés de ce même appareillage. Les orbites des deux satellites ont une inclinaison semblable, à environ 2,6 degrés de distance dans l'ascension droite des noeuds ascendants. On a choisi ces orbites pour qu'une zone située sur l'équateur puisse être couverte par les détecteurs à orientation verticale du premier satellite et par les détecteurs inclinés à 30 degrés environ du nadir du deuxième satellite. Grâce à cette configuration, un lieu précis peut être reproduit sur six images, soit trois par satellite.

Les orbites du système seront les plus basses possibles, héliosynchrones, peut-être à dérive minimale, presque circulaires et polaires. Des récepteurs SPG embarqués détermineront la position des satellites, tandis qu'un ou plusieurs détecteurs stellaires établiront l'attitude des détecteurs.

On se sert de la forme générale de l'équation de colinéarité dans laquelle une fonction linéaire et des fonctions d'ordre supérieur s'intercalent entre les images d'orientation pour mettre en rapport les coordonnées d'image, les paramètres de détecteur et les coordonnées-terrain.

On mesure la précision avec laquelle les positions de points sont déterminées en procédant à une analyse des variances-covariances. Dans le cas du système à un satellite, on obtient des résultats de 1,2, 2,0 et 3,9 mètres respectivement dans les directions X, Y et Z, lorsque la précision est de 1/10 pixel pour la mesure des images, de 5 cm pour le positionnement SPG et de 2 secondes d'arc pour les détecteurs stellaires. Avec les mêmes paramètres et six images, le système à deux satellites offre une précision de 1,0, 1,0 et 2,0 dans les directions X, Y et Z. À la fin de cette étude, on proposera un système satellitaire en temps réel dont la configuration permettra d'obtenir la meilleure précision qui soit et un degré élevé de fiabilité, sans avoir recours à des points de contrôle sur le terrain.

ABSTRACT

DEMONSTRATION OF SELECTED ASPECTS OF A UTILITY MAPPING SYSTEM (UMS)

by
Dean Merchant, Topo Photo, Inc.
Robert Tudhope, Ontario Hydro

Mapping power distribution lines presents a unique photogrammetric challenge. The homogeneous character of a power cable precludes the identification of conjugate imagery in the photo-base direction when photography is flown in the direction of the power line. By locating small, digital array cameras in the wing tips of low flying aircraft, the error in image identification occurs in the cross-base direction (y-parallax) and does not present a problem in computing conjugate ray intersections for measurement of elevation and horizontal position normal to the direction of the cable. A program of research was sponsored by Ontario Hydro of Toronto to test this "cross-base" concept for power line surveys. Results indicate that the cross-base concept is feasible and can meet required spatial accuracies provided care is taken in providing strong geometry and calibrating the photogrammetric system in the air. Requirements by utilities for spatial accuracy on cables, towers and insulator attach points are described. Results obtained from flight tests are presented.

DÉMONSTRATION DE CERTAINS ASPECTS D'UN SYSTÈME DE CARTOGRAPHIE POUR LES ENTREPRISES DE SERVICE PUBLIC

RÉSUMÉ

La cartographie des lignes de transport d'électricité représente un défi unique pour la photogrammétrie. Le caractère homogène des câbles électriques empêche l'identification d'images conjuguées dans la base de prises de vue lorsqu'on photographie en se déplaçant parallèlement aux lignes. En disposant une série de petites caméras numériques à l'extrémité des ailes d'un aéronef volant à basse altitude, l'erreur d'identification d'image se produit dans le sens de la base transversale (parallaxe en y) et ne présente pas de problème pour le calcul des intersections de rayons conjugués permettant de mesurer l'altitude et la position horizontale normales par rapport à la direction du câble. Ontario Hydro, à Toronto, a parrainé un programme de recherche visant à mettre à l'épreuve ce concept de «base transversale» pour les levés de lignes électriques. D'après les résultats obtenus, le concept serait viable et permettrait d'atteindre la précision spatiale voulue, à condition de satisfaire à des exigences strictes en matière de géométrie et d'étalonnage du système photogrammétrique pendant les vols. On décrit les exigences des entreprises de service public quant à la précision spatiale des câbles, des pylônes et des points de fixation des isolateurs. Les résultats obtenus lors des essais en vol sont présentés.

A System for Underwater Stereo Video Image Processing and Its Application in Fisheries

R. Li[1], T. Curran[2], R. Smith[2], H. Li[1]

[1]Department of Geomatics Engineering, The University of Calgary, Calgary, AB T2N 1N4
[2]Institute of Ocean Sciences, Sidney, B.C. V8L 4B2

Abstract

Underwater video images are mostly used for visualizing objects in the underwater environment. Quantitative analysis, such as measuring objects and reconstruction of 3D object shapes, still needs to be researched. In this paper, a system for underwater stereo video image preprocessing is presented, which includes modules of Graphical User Interface (GUI), preprocessing, camera calibration, photogrammetric processing and graphic display. Some geometrical, optical and electrical properties of CCD-camera and images in the underwater environment are investigated, for instance, lens distortion, noise elimination/reduction, ray bending etc. Based on these, a photogrammetric model can be established for determining object space coordinates from measured image coordinates. This allows to calculate positions and shapes in the 3D object space as long as the objects to be measured can be seen in stereo images. The GUI makes the modules involved transparent to users and provides a convenient, efficient and user-friendly environment for object-oriented measuring procedures. Results of a water tank test for the camera calibration and fisheries applications are presented. Other potential applications of the system could be precise seafloor mapping, underwater target tracing, object monitoring and others.

Système de traitement des images vidéo stéréoscopiques sous l'eau et applications en halieutique

Résumé

Les images vidéo prises sous l'eau servent surtout à visualiser les objets présents dans le milieu marin. L'analyse quantitative de ces objets, notamment leur mesure et leur reconstruction en trois dimensions, demande encore beaucoup de recherche. Dans ce document, on présente un système de prétraitement des images vidéo stéréoscopiques sous l'eau; ce système comporte des modules d'interface utilisateur graphique (IUG), de prétraitement, d'étalonnage de la chambre, de traitement photogrammétrique et d'affichage graphique. On y aborde aussi certaines des propriétés géométriques, optiques et électriques des caméras à dispositif de couplage de charge servant à produire des images en milieu marin, par exemple, la distorsion de l'objectif, la réduction ou l'élimination du bruit et la courbure des rayons. À l'aide de ces paramètres, il est possible de construire un modèle photogrammétrique permettant de déterminer les coordonnées spatiales des objets à partir des coordonnées mesurées sur l'image et de calculer ensuite la position et la forme des espaces-objets en trois dimensions, dans la mesure où les objets à mesurer sont visibles sur les images stéréoscopiques. Grâce à l'IUG, les modules fonctionnent en mode transparent et offrent aux utilisateurs un environnement d'une grande commodité, efficace et convivial pour les opérations de mesure orientées objet. On y présente par ailleurs les résultats d'un essai en piscine d'étalonnage de chambre ainsi que les applications en halieutique. Le système pourrait, entre autres applications, permettre une cartographie précise du plancher océanique, la poursuite de cibles marines et la surveillance d'objets.

SELECTION OF AN OPTIMUM STRUCTURED METHODOLOGY FOR DEVELOPING AN INFORMATION SYSTEM; A GENERAL MODEL.

C.M. Paresi, M.M. Radwan, O.U. Essien

International Institute for Aerospace Survey and Earth Sciences (ITC)
P.O. Box 6, 7500 AA Enschede, The Netherlands

KEY WORDS: Information Systems, Structured Development Methodologies, Optimum Methodology, Evaluation Model.

ABSTRACT

Where the overall intent is a realistic and organised control of the System Development process, with a resultant effective Information System, then it becomes imperative that the Methodology chosen for the development of the System, should be optimum for that Object System. The paper shortly reviews the characteristics of Structured Development Methodologies. It then focuses on the determination of a general evaluation model criteria for selecting an optimum Structured Development Methodology applicable to a large number of problem domains. An overview of the model is first presented, followed by an explanation of the steps in the process of evaluation.

INTRODUCTION

In the development of Information Systems which could be Administrative, Technical, or Managerial, Structured Development Methodologies have been used to create systems that give an effective communication between the system users, the machines that support the systems operations, as well as the human factors in information exchange. The bottom line is an efficient and organised control of the system development process.

The number of these Methodologies have of recent increased considerably in the Information Science/ Technology scene. Some look similar in many respects, some use techniques that overlap in concepts, the advantages or otherwise of some are more apparent in some types of Systems than the others, and some lay more emphasis on some aspects of Information Systems development.

Therefore, a research was initiated on the determination of a general *evaluation model criteria* for selecting an Optimum Structured Development Methodology applicable to a large number of problem domains.

The philosophy of the model is based on the fact that it is unlikely at the moment, to have a single methodology that can satisfy all the phases and stages during system development. Therefore for a particular Object System, development phases/stages are identified according to their relative importance. The methodologies are then examined for how well they provide for the individual phases/stages. Based on the peculiar circumstances of the Object System under study, several contingencies/criteria can be determined and embedded in a general framework through which the methodologies would be evaluated.

The model makes provisions for levels of detail to be examined in the Object System, the levels of decomposition to which phases/stages of the Object System can be reasonably subjected, and also a means of reducing subjectivity inherent in this type of evaluation. It is possible to carry out the evaluation either for each phase, or the entire phases put together. It can also be decided using a contingency sub-model, whether the development strategy should be the use of only prototypes, specifying, or a mixture of both. The model recognises the influences from issues collateral to the Object System.

WHY STRUCTURED DEVELOPMENT METHODOLOGIES?

Systems Development is a non-trivial task; it costs time and money because of the complex nature of modern Information Systems and therefore needs a clear step by step approach in the process of development. This is the concept behind Methodologies. It defines the pieces or components, phases and activities that one finds in a typical System Development project; and also the interface between those components [Burch,1979]. The most important factor for the need for Methodologies is the limitation of the human mind to perceive and retain all information it requires and to act on it promptly [Dippel, 1969].

A Methodology should have as general objectives, to : Analyse the complexity of problems and simplify them; achieve unity of System Architecture; establish and improve interaction between users and System Development Team; enable efficient parallel development of Sub_Systems; achieve sound data analysis and administration through a well defined procedural framework [Elving & Kirchoff,1991].

A Methodology should specify:
How a project is to be broken down into phases, activities for each phase, outputs for each phase, when activities should be executed, constraints to be applied, support tools to be utilized, and how projects have to be managed. Techniques (which can be diagrams, tables etc.) should be recommended at various stages for carrying out particular activities recommended by a Methodology.

Sélection d'une méthode structurée optimale pour le développement d'un Système d'information : un modèle généralisé

RÉSUMÉ

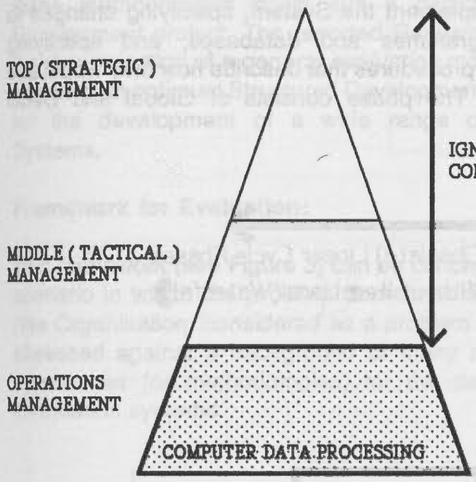
Il est impératif que la méthode utilisée pour le développement d'un Système d'Information soit parfaitement adaptée au domaine particulier du système en question, afin que le Système d'Information réponde d'une manière effective aux besoins des utilisateurs.

La communication présente tout d'abord les raisons d'être et les caractéristiques des Méthodes de Développement Structurées de Systèmes d'Information. Elle considère ensuite, le développement d'un *modèle d'évaluation généralisé*, servant à sélectionner la Méthode de Développement Structurée la mieux adaptée, pour un grand nombre de domaines. Le modèle et la procédure d'évaluation sont présentés et discutés.

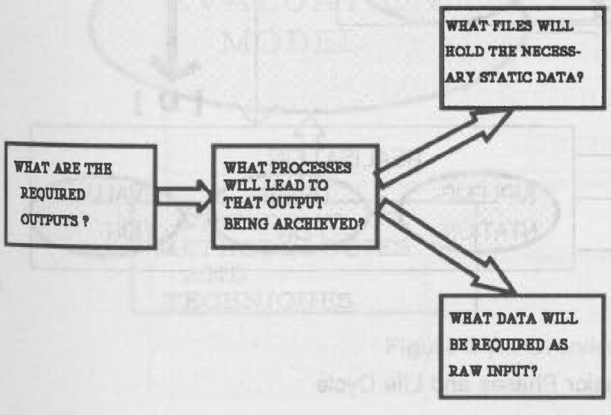
Benefits from Methodologies should include high quality products which are easy to maintain and upgrade, better control during the design process, reduction time to completion and lower costs of developments.

Unstructured Methodologies

Generally referred to as "Conventional Approach" in most texts, they where prevalent in the 1970s. The philosophy behind these Methodologies is visualised as shown in Figure 1.



(a). Failure to meet all Organisational needs.



(b). Output Driven Design.

Figure 1 A Window on Unstructured Methodologies (Adapted from Avison & Fitzgerald)

The Unstructured Methodologies were unsatisfactory due to:

Failure to meet the needs of management, instability and inflexibility, user dissatisfaction, problem of documentation, incomplete Systems, little or no control over the quality of the design, and lack of an objective criteria with which to judge whether the critical activities of analysis and design were carried out properly [Yourdon,1988].
These create concern for:

Low productivity, low quality, long development cycles, increasing costs of development and maintenance of Systems, and thus create the need for Structured Methodologies [Paresi,1991].

The concern for more productive methods for the development of information systems has therefore tremendously increased during the last two decades. A wealth of research has addressed the problem of developing efficient and effective Information Systems. Four approaches to development methodologies have emerged from research and experience; these are the formal methods, the structured methods, the soft methods, and the socio-technical methods. In practice, any development is a combination of these different approaches; only Structured Development Methodologies will be considered here.

Structured Systems Development Methodologies

Necessity and Basic Features:

Given the generic objectives of Methodologies as stated earlier, it is obvious that conventional approaches fail particularly when Systems become more complex. Structured Methodologies seek to address these failures as outlined above with particular consideration to the reduction of the cost of rectifying changes in user requirements. Such changes would remain minimal if these requirements are considered extensively during the analysis phase [Cutts, 1991]. They also seek to mirror organisational patterns faithfully during implementation [Ward,1984]. Some of the fundamental features of Structured Methodologies include:

(1) A Top-Down approach with a strict philosophy of separating the conceptual System from the physical System. Much effort is put developing the conceptual system during which the analyst gets an overall grasp of the System, breaks the picture down in manageable details, with a productive use of time.
(2) They avoid the duplication of efforts and waste of resources, thereby reducing the volume of work and speed up the overall System Development process.
(3) Much of the documentation, which is in the form of diagrams and systems encyclopedia (or data dictionary), are better understood by non-technical users, and clear thinking is encouraged.
(4) By encouraging standardized methods, another designer can easily assimilate the steps, and System maintenance becomes less cumbersome.
(5) The Techniques and Tools are not 'doctrinal' or inflexible and force the analyst to ask questions to users, and of himself, thereby minimising wrong assumptions during design.
(6) Most parts of the analysis and design can be developed, maintained and held on computer systems (CASE tools), easing the work of documentation.

Phases and Scope of Structured Methodologies:

The modification and development of Information Systems in present day organisations is a dynamic and sometime controversial activity. The steps recommended by Structured Methodologies in building Information Systems consists of a set of iterative activities referred to as the System or Project Life Cycle. These activities are grouped in Phases.

The Major Phases are: (see also Figure 2)

[A] - Systems Strategy and Planning
[B] - Systems Analysis
[C] - Systems Design
[D] - Systems Realisation

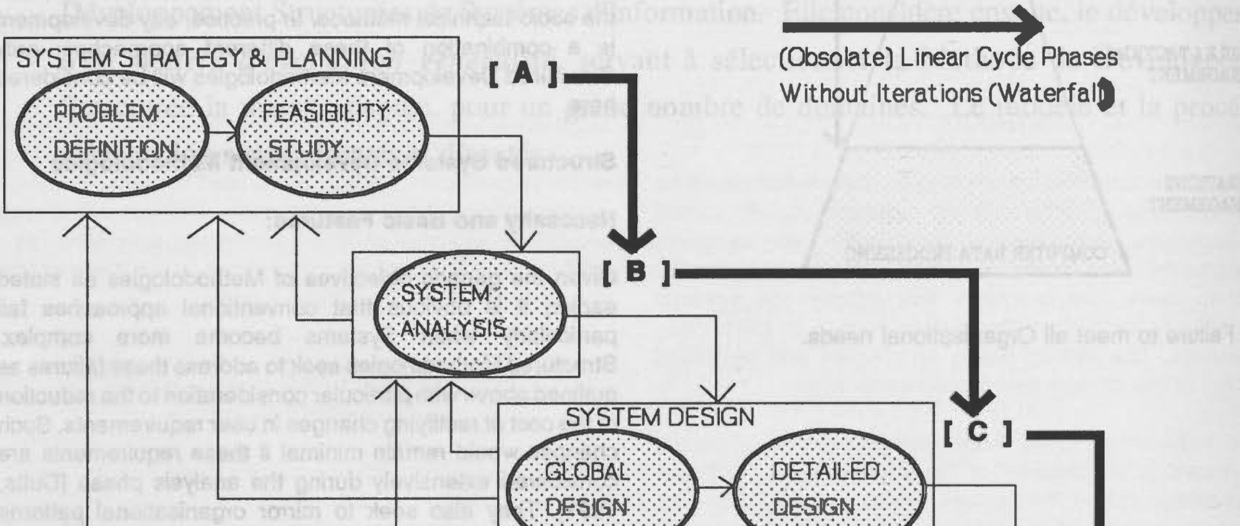


Figure 2 System Development major Phases and Life Cycle

[A] - System Strategy and Planning Phase:
Considered as the most important phase it has as objectives to define the problems to be solved using as terms of reference:

<u>project goals</u> : needs and important sub-system(s)
<u>project bounds</u> : System boundaries and relation between the organisational strategy and IT, and
<u>resource limits</u> : time, costs, personnel, infrastructure and facilities.

It consists mainly of a Definition Study and of a Feasibility Study.

[B] - System Analysis Phase:
As objective, this phase seeks to make a detailed appraisal of the *existing* System in the organisation, and to develop alternative conceptual solutions for the new System.

[C] - System Design Phase:
Objectives here include selecting engineering design needed to implement the System, specifying changes to existing programmes and databases, and specifying detailed user procedures that describe how they would use the System. The phase consists of Global and Detail design.

[D] - Realisation Phase:
This can be seen in two perspectives: Implementation and Installation; and Operation/Maintenance. As objectives, it aims to turn to fruition, all that had been put down on paper in previous phases and to ascertain how worthy the whole project was.

A GENERAL EVALUATION MODEL CRITERIA FOR SELECTING AN OPTIMUM STRUCTURED DEVELOPMENT METHODOLOGY

A lot of Structured Development Methodologies exist in the Information Science/ Technology scene. Some look similar in many respects, some use techniques that overlap in concepts, the advantages or otherwise of some are more apparent in some types of Systems than the others, and some lay more emphasis on some aspects of Information Systems development. This state of affairs leaves the Systems Analyst and Designer in a quandary over which of these methodologies could suit a particular Systems Development project. The reported Research focuses on the determination of a general *evaluation model criteria* for selecting an optimum Structured Development Methodology for the development of a wide range of Information Systems.

Framework for Evaluation:

This framework (see Figure 3) can be conceptualised as a scenario in which the problem domain or *Object System* (the Organisation, considered as a problem area) is to be assessed against a background of many recommended approaches (or methodologies) for the development of information systems.

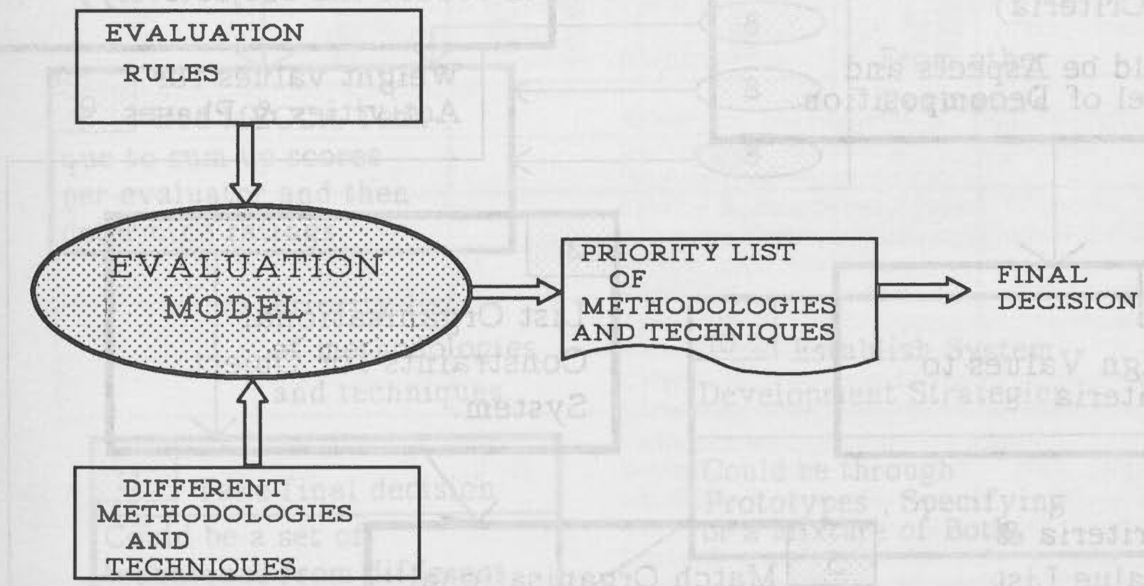


Figure 3 An Overview of the *Evaluation Process*

The *Evaluation Rules* should include the following guidelines (based on the nature of the problem area for the Object System):

- The composition of an evaluation committee
- Aspects (perspectives) of methodologies to be evaluated
- Level of decomposition (as provided for by the methodology) to be considered or seen as desirable.
- The level of details to be evaluated per methodology
- How the extent of subjectivity in evaluation (particularly weighting) can be reduced (could be through iterations using the Delphi Technique of quantitative studies)
- How decisions can be taken on very closely scored methodologies (could be through increasing the levels of details evaluated)

The *Different Methodologies* would be introduced into the evaluation framework
(matrices or tables) determined in the evaluation model.

The *Evaluation Model* would simply be a model of selection processes, resulting in a priority list of methodologies,and also in a final decision. This model is the core of the scenario and is blown further as shown in Figure 4.

Further Explanation of the Framework

♦ STEP 0: This involves outlining of the principal phases through which the development process for the Object System would (or should) have to go through. In the presented example of Topographic Database Systems, the phases of development are [Radwan, undated] Design, Implementation and Operation.

♦ STEP 1: Here the functions and/or activities within the phases defined above are listed.

♦ STEP 2: The importance of each phase and function/ activity is weighted. This is done independent of any considerations for any methodology. It would have more to do with the peculiar circumstances of our Object System, and each evaluator would state reasons why a particular phase, function/activity is given more importance relative to the rest. Viewed as an individual judgement, weighting can be highly subjective because the phases are not mutually exclusive and the evaluators have different perceptions and experiences. Through iterations, the evaluators can re-evaluate their weights (where discrepancies are unacceptable) and move towards an acceptable mean

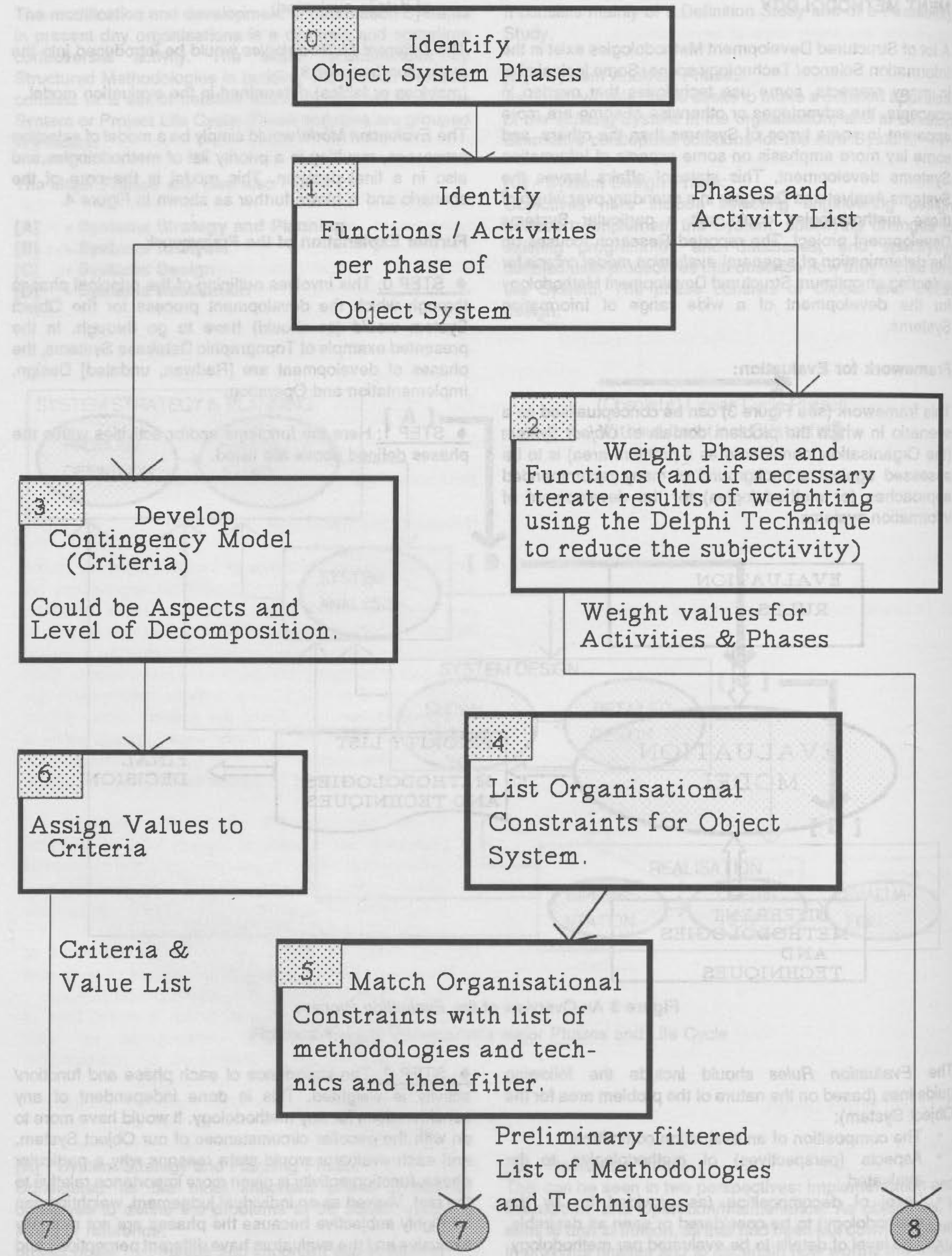


Figure 4 Scenario of the Evaluation Model (*Continued*).

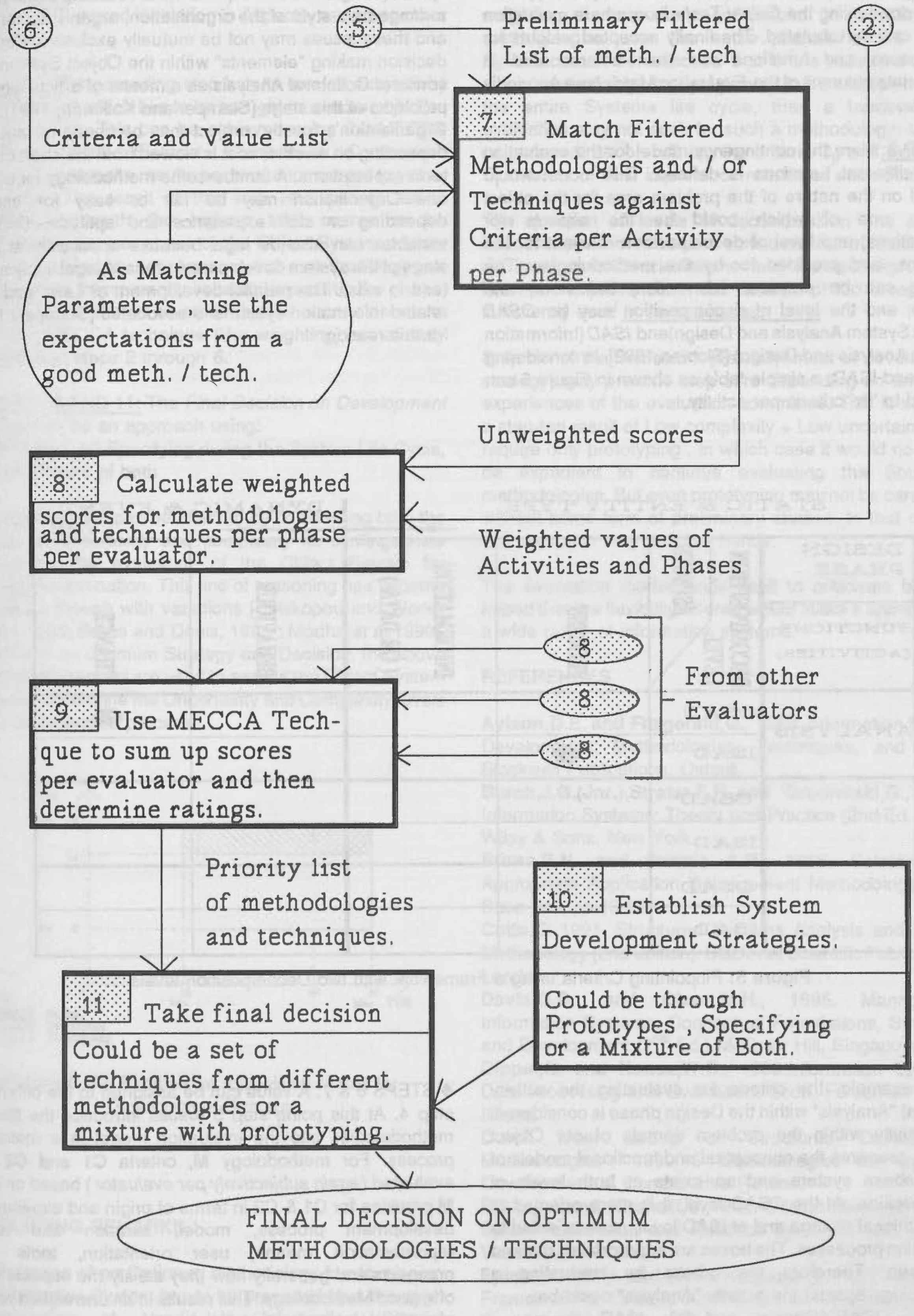


Figure 4 Scenario of the Evaluation Model.

weight (which implies a better measure of objectivity). This can be done using the *Delphi Technique*, where evaluation results can be tabulated. The finally accepted weights for the phases and functions can then be recorded in appropriate columns of the *Evaluation Matrix* (see Appendix 1).

♦ STEP 3: Here the contingency model for the evaluation of the different functions is defined. This model would depend on the nature of the problem area for the object system, one of which could be the aspects (or perspectives) and level of decomposition relevant for the functions, and provided for by the methodologies. The aspects can be *process, information, behaviour, and change*; and the level of decomposition may be *OSAD* (Object System Analysis and Design) and *ISAD* (Information System Analysis and Design) [Slooten, 1992]. In considering OSAD and ISAD, a simple table as shown in Figure 5 can be used to 'fix' criteria per activity.

specific. Decisions can depend on far reaching concerns as management style of the organisation, organisational size; and these issues may not be mutually exclusive. Perhaps decision making "elements" within the Object System may consider Collateral Analysis as a means of articulating the problems at this stage [Stamper and Kolkman, 1991]. One Organisation's "costly" project may be cheap for another, depending on whether cost is viewed from the short or long term expectations. A cumbersome methodology for staff of one Organisation may be fair or easy for another depending on staff experience and aptitude. Decision variables can also be legal because apparently at each stage of the system development process, legal implications tend to exist. The parallel development of Laws and their related information systems is advocated [Stamper, 1992] for this reason.

STATIC & ENTITY TYPE ← → DYNAMIC & EVENT TYPE

DESIGN PHASE / FUNCTIONS (ACTIVITIES)	DECOMPOSITION LEVEL / ASPECTS	PROCESS	INFORMATION (OR DATA)	BEHAVIOUR	CHANGE
ANALYSIS	OSAD				●
	ISAD	●			
	OSAD				
	ISAD				
	OSAD				
	ISAD				

Figure 5: Pinpointing Criteria using a Framework with two Decomposition levels

As an example, the criteria for evaluating the activity (function) "Analysis" within the Design phase is considered. This activity within the problem domain of our Object System, prepares the conceptual and functional models of the database system and so exists at both levels of decomposition. At the OSAD level, it is more oriented to Organisational change and at ISAD level, more oriented to information processes. The boxes are appropriately marked as shown. Therefore the criteria for evaluating a methodology against the activity "Analysis" could be:

C1 = OSAD / Change* and *C2 = ISAD / Process.

♦ STEPS 4 AND 5: These mark a decision process tempered by the global issues inherent in the Object System. These could be Institutional, Economical and Technical and are usually Organisation and System

♦ STEPS 6 & 7: A value can be assigned to the criteria in step 4. At this point, step 5 results introduce the filtered methodologies into the evaluation model in a matching process. For methodology **M**, criteria **C1** and **C2** are evaluated (*again subjectively per evaluator*) based on how **M** provides for C1 & C2 in terms of origin and experience, development process, model, iteration and tests, representation means, user orientation, tools and prospects, and generally how they satisfy the expectations of a good Methodology. This results in an unweighed rating of methodologies for that activity.

♦ STEP 8: Here the ratings are weighted, first per function, then per phase, and then a simple sum of weights is determined for all phases. This would conclude the preliminary stage of the process of selection per methodology per evaluator based on the contingency model (OSAD,ISAD viewed in relation to Aspects, as well as other criteria).

♦ STEP 9: The preceding step has its level of subjectivity (per evaluator) which has to be brought to a *more objective* level of choice. This can be achieved using the MECCA technique (*Multi Element Component Comparison Analysis*). Composite results per evaluator are put together in a table. The result of this step is a priority list of methodologies with the highest total scores to be considered either per phase of the Object System or as a whole, but preferably per phase; so that a methodology can be seen to be transparently suitable for a particular phase, or otherwise. It should be noted that other considerations within the MECCA technique (like weighting) are already embedded in steps 2 through 6.

♦ STEPS 10 AND 11: The *Final Decision on Development Strategy* can be an approach using:
(i) Prototypes, (ii) Specifying during the System Life Cycle, or (iii) A Mixture of both

The determination can be effected by considering both the degrees of *uncertainty* and *complexity* as contingencies relevant for the evaluation of the Object System for Strategy determination. This line of reasoning has become acceptable though with variations [Episkopou and Wood-Harper, 1986; Burns and Denis, 1985 ; Modha et al,1990]. To arrive at an optimum Strategy and Decision, the above stated *contingencies* are used to assess the Object System in order to determine the Uncertainty and Complexity levels of the development process.

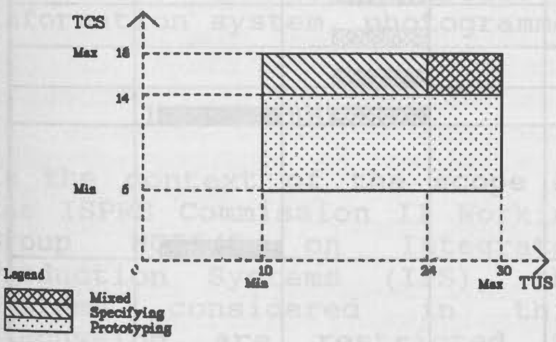


Figure 6 An Example of a Decision Matrix for Development Strategy

CONCLUDING REMARKS

The Selection of an Optimum methodology for developing any Information System has to be preceded by a thorough understanding of trends in Information System Development, the nature of the Object System, as well as a good reasoned analysis of any criteria and contingencies necessary for use in the selection process.

The criteria score can directly be used without the phases weighting if one does not intend to evaluate the methodologies over the entire phases of the Object System. It can be reasoned that if it is accepted that no single methodology has yet proven to be satisfactory over the entire Systems life cycle, then a framework for evaluation may not seek for such a methodology; rather it can concentrate on refining the criteria C_i, and examining C_i in terms of their importance relative to each other.

The question of levels of decomposition and aspects become apparent when one considers that methodologies do have specific orientations and the techniques they recommend prove either too detailed or too broad to be sufficiently used in some phases.

Step ten of the framework (Establishment of development strategy) may precede step zero depending on any other experiences of the evaluation committee. This is so since a step-ten result of Low complexity + Low uncertainty may require only prototyping , in which case it would no longer be expedient to continue evaluating the Structured methodologies. But even prototyping may not be carried out without some form of preliminary studies. In that case, a *soft* approach may become handy.

The evaluation model lends itself to criticisms but it is hoped that the flexibility offered would make it applicable to a wide range of information systems.

REFERENCES

Avison,D.E. and Fitzgerald,G., 1988. Information System Development: Methodologies, Techniques, and Tools. Blackwell Publications, Oxford.

Burch,J.G.(Jnr.),Strater,F.R.,and Grudnitski,G., 1979. Information Systems: Theory and Practice (2nd.Ed.). John Wiley & Sons, New York.

Burns,R.N. and Dennis A.R., 1985. Selecting the Appropriate Application Development Methodology. Data Base Fall, pp.19-23.

Cutts,G.,1991. Structured Systems Analysis and Design Methodology (2nd Edition). Blackwell Scientific Publications, London.

Davis,G.B., and Olson,M.H., 1985. Management Information Systems: Conceptual Foundations, Structure, and Development.(2nd Ed.) McGraw Hill, Singapore.

Dippel,G. and House,W.C., 1969.Information Systems: Data Processing and Evaluation. Scott, Foresman & Co., Illinois.

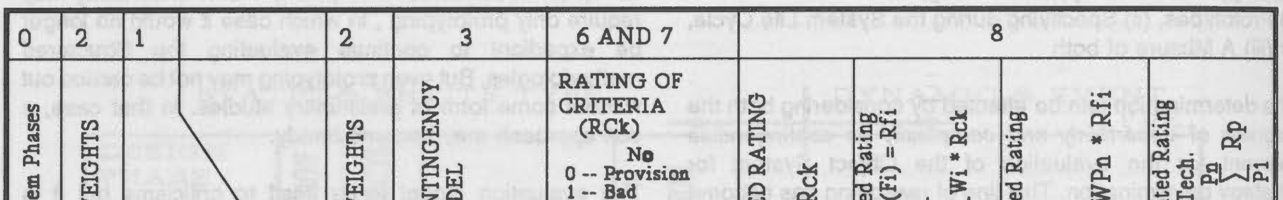
Dutch User Group of Structured Development Methodologies", 1990.16 Methodologies for Systems Development: A Comparative Review.

Elfving,A.and Kirchoff,U.,1991. Design Methodology for Space Automation and Robotics Systems. ESA Journal Vol.15, pp.149-164.

Episkopou,D.M. and Wood-Harper,A.T.,1986. Towards a Framework to Choose Appropriate IS Approaches. The Computer Journal, vol. 29, No. 3, pp.222-228.

Essien,O.U.,1992. Selection of an Optimum Structured Methodology for the Development of Topographic Database Systems. MSc. Thesis, ITC.

Martin,J. and McClure,C.,1985. Diagramming Techniques for Analysts and Programmers. Prentice-Hall Inc. Englewood Cliffs, New Jersey.
McRae,S.D.,1989. GIS Design and the Questions Users should be asking. GIS/LIS Proceedings, Orlando, Florida, November 26th-30th, vol.2,pp.528-537.
Mittra,S.S.,1988. Structured Techniques of System Analysis, Design and Implementation. John Wiley and Sons, New York.
Modha,J., Gwinnett,A. and Bruce,M.,1990. A Review of Information Systems Development Methodology (ISDM) Selection Techniques. OMEGA International Journal of Management Science,vol. 18,No.5,pp.473-490.
Naumann,J. and Palvia,S.,1982. A Selection Model for Systems Development Tools. MIS Quarterly / March, pp.39-47.
Olle,T.W. et al.,1991. Information Systems Methodologies: A Framework for Understanding (2nd.Ed.). Addison-Wesley, England.
Paresi,C.,1991. Introduction to Information System Development. ITC Lecture Notes Series.
Radwan,M.M.(Undated). Digital Mapping and Topographic Databases. ITC Lecture notes.
Slooten,K. van.,1992. A Multi-Level Approach for the Analysis and Design Process. Proceedings MOACSI, Nonkes, Sept.
Stamper,R. and Kolkman,M.,1991. Problem Articulation: A Sharp-Edged Soft Systems Approach.Journal of Applied Information Systems,vol.18,pp.69-76.
Stamper,R.,1992. The Parallel Development of Laws and their related Information Systems. University of Twente, The Netherlands.
Ward,P.T.,1984. Systems Development Without Pain: A User's Guide to Modelling Organisational Patterns. Yourdon Press, New Jersey.
Yourdon,E.,1988. Managing the System Life Cycle. (2nd Edition) Yourdon Press, New Jersey.



APPENDIX 1: The Evaluation Matrix per Methodology/Technique per Evaluator *(for steps 0 through 8)*. Steps 4 and 5 are under Organisational Constraints.

INSTITUTIONALIZATION OF INTEGRATED PRODUCTION SYSTEMS FOR SPATIAL DATA

ROOP C. MALHOTRA, Ph.D

National Geodetic Survey Division
Coast and Geodetic Survey
National Ocean Service
National Oceanic and Atmospheric Administration
Silver Spring, MD 20910 U.S.A.

ABSTRACT

The ultimate goal of Integrated Production Systems (IPS) for spatial data is to serve the needs of the user community in a cost effective and efficient manner. The problems facing the process of institutionalization of such systems must be addressed in order to enhance their usage. Besides the merits of the IPS, which are based on the advanced technologies involved in the their development, the driving force behind their usage will be their capacity to integrate into widely used geographic information systems. A greater understanding of the issues in this context will help in avoiding major pitfalls and lead to wider acceptance of IPS. In this paper, some of the major institutional issues are identified, discussed, and recommendations made.

KEY WORDS: Integrated production systems, spatial data, geographic information system, photogrammetry, remote sensing.

1. INTRODUCTION

In the context of the scope of the ISPRS Commission II Working Group WGII/5 on Integrated Production Systems (IPS), the systems considered in this discussion are restricted to those which generate geo-referenced data products from spatial data using technologies of photogrammetry and remote sensing. IPS makes use of photogrammetric source data to produce topographic maps, ortho-photo images and maps, digital terrain models and the like. Likewise, IPS uses remotely sensed source data from the Image Analysis System (IAS) to produce thematic maps. Interface of IPS with Geographic Information System (GIS) enhances the realm of applications and products of IPS. This is due to the inherent capability of GIS to perform analysis of several layers of spatial data to produce GIS-derived spatial information for specific user applications. Data bases for spatial data and its attributes are an integral part of the GIS for analysis and its subsequent use by the IPS.

2. BACKGROUND

2.1 Technologies

Photogrammetry and remote sensing technologies provide geo-referenced source data for IPS. The IAS manipulates remotely sensed digital image data as a source data for the IPS. Photogrammetric plotters collect map compilation data which is yet another source of spatial data for the IPS. GIS technology facilitates the storage of and access to many types of data. With the user demand for spatial data products on the increase, integration of different data sources has become essential in the design of IPS. Spatial data IPS when interfaced with GIS combines cartographic, raster image-based, and tabular ancillary data for many useful products. Furthermore, incorporation of temporal data into a spatial database increases the realm of IPS products.

Technological advances in the IPS may be measured in terms of advances in the related technologies of GIS, satellite remote sensing, IAS, photogrammetry, multimedia computers and databases. The current trend in computer industry indicates that multimedia computer platforms will soon become general tools for information technology. These multimedia computer plat- and provide capability to integrate data, text, images, video and voice for potential applications. Incorporation of such a multimedia in IPS will enable the use, manipulation and presentation of both static and dynamic spatial data.

2.2 Institutional issues

The process of technology acceptance in any field is complex, and research results often require several years to be adopted in practical applications or operations. Although the technologies involved in the IPS are well established, their integration in IPS has not been fully achieved. This may be attributed to impediments imposed by certain institutional issues rather than technical constraints. The ISPRS WG and other interested groups must task themselves to identify, and recommend ways and means to remove such impediments in the way of integration of technologies and widespread use of IPS in various applications.

3. INSTITUTIONALIZATION OF IPS

The National Center for Geographic Information and Analysis (NCGIA) was formed by the U.S National Science Foundation to remove impediments to the broader application of GIS and Geographic Analysis. This is a problem similar to the one that the WG will face in the process of institutionalization of IPS - the identification of institutional issues and finding their solutions for widespread use of IPS.

NCGIA came up with six institutional issues affecting the use of integrated remote sensing and GIS technologies. The issues are: (1) data availability, (2) data marketing and costs, (3) equipment availability and costs, (4) standards and practices, (5) education and training, and (6) organizational infrastructure (Lauer, et al. 1991)). A parallelism is apparent in the use of IPS for spatial data applications.

The utility of an IPS is enhanced manyfold when interfaced with a GIS such that the problem of data availability is minimized provided the GIS has the desired data. Moreover, the standard products supported by a GIS must meet the needs of the users. Many of the problems related to the data availability have been recognized. Data availability issue is reflected in the 1990 Office of Management and Budget (OMB) Circular No. A-16 "to improve coordination of surveying, mapping, and related spatial data activities" with the intent to develop a national digital spatial information resource, avoid duplication, and encourage sharing of data amongst various institutions.

Restrictions and cost of reproduction of spatial data by parent institutions must be minimized for a wider usage of the data. An example of such a gesture is that National Oceanic and Atmospheric Administration and Earth Observation Satellite Company announced that the Landsat data more than 2-years old could be purchased for the cost of reproduction (Lauer, et al. 1991).

Consideration must be given to other institutional issues related to budgetary constraints, policy matters, proprietary interests, training, and potential data products for users. Acquisition of IPS hardware and software is limited by the available budget. For this reason, basic understanding of system development methodologies will help develop IPS specifically suited for user needs. Policy must be established regarding systems development. For example, standard configuration for IPS may not be conducive to the evolutionary process for new technological breakthroughs. On the other hand, interface technology encourages such a development. It is also essential to maintain a highly trained support staff to keep the IPS operational. Public domain software may be cheaper to obtain but in the long run may prove to be more costly when applied to operational problem solving.

4. RECOMMENDATIONS

A viable infrastructure must be established for IPS by providing proper management, equipment (hardware and software), data resources (GIS and source data) and training for both current and future activities or products.

In order to benefit the user community, the existing IPS and their products must be widely publicized. There may be reluctance on the part of an institution to undertake an extra step of publicizing IPS due to budgetary constraints unless incentives are provided for doing so. Also, in order to be marketed, an existing IPS must fully demonstrate its capabilities as well as be operational for the targeted products.

On the other hand, an organization interested in establishing a suitable IPS for its specific application must thoroughly study the existing IPS from all points of view, technological and institutional, in the context of the products desired. An IPS could be characterized by the technologies used, functionality, configuration (hardware and software), interfaces, sources for spatial data, training requirements, costs involved, products supported, and so forth. A well characterized IPS by its manufacturer, vendor or developer will enlighten the users for a better understanding of the existing systems. Users are then able to make an intelligent choice in putting together or integrating an efficient IPS for their specific needs.

It is, therefore, recommended that both the present users and manufacturers of IPS work closely to provide the future users with sufficiently detailed information on the existing IPS and their products. Also, institutional issues, as pointed out by the NCGIA, established by the NSF for broader application of GIS and Geographic Analysis, must be studied closely in relation to the IPS by such organizations as the ISPRS WG for promoting widespread usage of the IPS.

5. REFERENCES

Lauer, Donald T., Estes, John E., Jensen John E., and Greenlee, David D., 1991. Institutional Issues Affecting the Integration of Remotely Sensed Data and Geographic Information Systems. Journal of the American Society of Photoogrammetry and Remote Sensing, Photogrammetric Engineering and Remote Sensing, 6:647-654.

OMB, 1990. Coordination of Surveying, Mapping, and Related Spatial Data Activities, Office of Management and Budget, Circular A-16 (Revised), 8 p.

Institutionnalisation des systèmes intégrés de production (IPS) pour les données spatiales

Résumé

La raison d'être des systèmes intégrés de production (IPS) de données spatiales est de répondre aux besoins des milieux d'utilisation de manière économique et efficace. Les problèmes que pose le processus d'institutionnalisation de ces systèmes doivent être réglés pour que ceux-ci puissent être exploités à leur pleine possibilité. Outre les qualités intrinsèques des IPS, qui ont été élaborés grâce aux plus récentes technologies, ces systèmes deviendront vite indispensables du fait qu'on peut facilement les intégrer aux systèmes actuels, largement diffusés, d'information géographique. Une bonne compréhension des problèmes inhérents à l'implantation des IPS permettra aux utilisateurs d'éviter les embûches et contribuera à accroître leur diffusion. Le présent document définit et explique quelques-unes des principales difficultés que pose l'institutionnalisation, puis énonce certaines recommandations.

The National Advanced Remote Sensing Applications Program, an Integrated System of Commercial Off-the-Shelf Components

Gary Vanderhaven
U.S.Geological Survey, Reston, VA. 22092
Clifford W. Greve
U.S.Geological Survey, Reston, VA. 22092
K. Eric Anderson
U.S.Geological Survey, Reston, VA. 22092

ISPRS Commission II Working Group 5

KEY WORDS: Integrated Production System, Photogrammetry, Remote Sensing, GIS

ABSTRACT

Integrated production systems have been defined as integrated systems that generate georeferenced products from spatial information obtained through photogrammetry and remote sensing. Providing an interface to a GIS enhances the system's range of applications by adding the GIS's capabilities for analyzing multiple layers of registered spatial information. The U. S. Geologic Survey, starting in 1992, began a search for such a system as part of its National Advanced Remote Sensing Applications Program initiative. The search centered on a system composed of standard commercially available products since resource constraints dictated that the system could not be custom developed. The search was successful and shows that such systems are indeed becoming commercially available. The capabilities that they possess are already sparking considerable interest among earth scientists and program administrators in a wide variety of civil agencies within the U. S. Government.

INTRODUCTION

*In 1992, U.S. Geological Survey (USGS) began defining a new capability to be known as the National Advanced Remote Sensing Applications Program (NARSAP). It would be designed to provide a facility, the systems, and the trained personnel to permit scientists and administrators within the civil agencies of the U. S. Government to use data derived from advanced remote sensors in environmental, resource, and safety areas. The system would combine the capabilities traditionally found separately in the areas of photogrammetry, image processing, and geographic information systems (GIS). In developing the NARSAP concept and in determining the hardware and software systems that would implement it, the USGS defined what has come to be called an integrated production system. It is a system that generates georeferenced products from spatial information obtained through photogrammetry and remote sensing. Providing an interface to a GIS enhances the system's range of applications by adding the GIS's capabilities for analyzing multiple layers of registered spatial information. Spatial and attribute data should be storable in data bases that are an integral part of the GIS. The NARSAP system incorporates traditional photogrammetric operations with image processing capabilities and a GIS in a digital environment. The system is fully capable of producing a wide variety of georeferenced products using the whole range of airborne and orbital imaging sensors. Its design is open ended and will allow the incorporation of new sensor models and capabilities in the future. The system to accomplish all of the above is completely off the shelf.

*Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

BACKGROUND

Approaches to the interactive use of remote sensing systems with GIS's differ significantly between the traditional mapping community and the environmental and land management communities. The mapping community has typically employed rigorous photogrammetric methods using metric photography obtained by calibrated camera systems to generate the equivalent of a map, whether in hard or soft copy. That information is then introduced into the GIS for further analysis. Commercially available systems have usually consisted of an interactive data capture capability for use with analog and analytical photogrammetric plotters. In most instances, data files from these data capturesystems are then exported into a GIS for further analysis and product finishing.

On the other hand, the environmental and resource management communities usually rely on multispectral remote sensing systems, such as Landsat and SPOT, to derive their data. Because of the favorable geometry inherent in many of these systems, adjustment to the GIS coordinate system has typically been accomplished with relatively simple polynomial warping models, based upon common control points, with no rigorous modeling of the geometry of the sensor system. Data from these remote sensing systems are usually extracted directly from the soft copy images, either by multispectral feature classification methods, by digitizing the vector outlines, or by combinations of both methods. The interface to the GIS is typically interactive on a common workstation with the image processing software.

After analyzing future remote sensing processing needs in the USGS and other U.S. Government civil agencies the USGS observed that the precieved boundaries between mapping and environmental remote sensing frequently overlapped in addressing current problems. Some of the more advanced remote sensing systems, such as SPOT, can look off nadir to a significant degree; therefore, the effects of relief displacement on the plotted position of features must be considered. This positioning requires mathematical rigor beyond the ability of the simple polynomial warping model. The National Aerial Photography Program (NAPP) provides a source for relatively high-resolution panchromatic and color infrared images that many users have found to be a useful complement to traditional remote sensing data. The prospect of the ready availability of digital orthophoto quadrangles (DOQ) derived from NAPP imagery increases the need to be able to exploit all of these sources concurrently in a common workstation environment.

The NARSAP is designed to provide an environment where personnel from various civil agencies, with help and guidance from USGS personnel, can use these advanced remote sensors when addressing their problems. These problems may include: disaster relief from such phenomena as volcanic eruptions, earthquakes, hurricanes, and floods; as well as scientific experiments on earth processes, land use and land cover determination, geologic structure, and stream mixing patterns; or monitoring of stream flood stages, erosion, wetlands loss, potential volcanic activity, land cover, and dam safety. This list is certainly not exhaustive.

Because virtually all of the civil agencies have ongoing programs that use remote sensing and GIS's, the NARSAP facility needed the capability to ingest information from these external systems and to transfer information back to these same external systems for further analysis. In particular, adherence to the Spatial Data Transfer Standard (FIPS 173) vector and raster profiles was desired, with implementation as soon as the profiles are operational. The NARSAP workstations were to be located near the Advanced Cartographic System (ACS), a primary component of the USGS's modernized map production system. Because the workstations would be required to exchange GIS and other data with the ACS workstations, there had be compatibility with that system.

The NARSAP system acquisition began with the intent to purchase a commercial off-the-shelf (COTS) system, specifying the system in the broadest possible functional terms so that the specifications could be met by vendors' standard product lines rather than customized products. Innovation on the part of respondents was welcomed. If a function required by the specification could be met by a product in an acceptable manner other than that defined by the specifications, the vendor was encouraged to address the alternative. It was preferred that the functional requirements be met by a truly COTS system rather than by special modifications to a system.

SYSTEM REQUIREMENTS

When the NARSAP specifications were being developed, three points were stressed:

1. The system specifications would have to be written in the broadest possible terms so that they could be met by standard product lines. No custom development was to be involved in the procurement.

2. The initial system configuration would have to provide basic processing capabilities for testing potential applications.

3. Because it is impossible to forecast what new capabilities will be developed or where they will come from, the initial system must be as open as possible to future enhancements.

A block diagram of a first draft is shown in figure 1. Its purpose was as much to help the USGS understand what was needed as it was to help prospective contractors. The final configuration was clearly left to the contractors to propose.

Workstations were to have stand alone capability. However, a Transmission Control Protocol/Internet

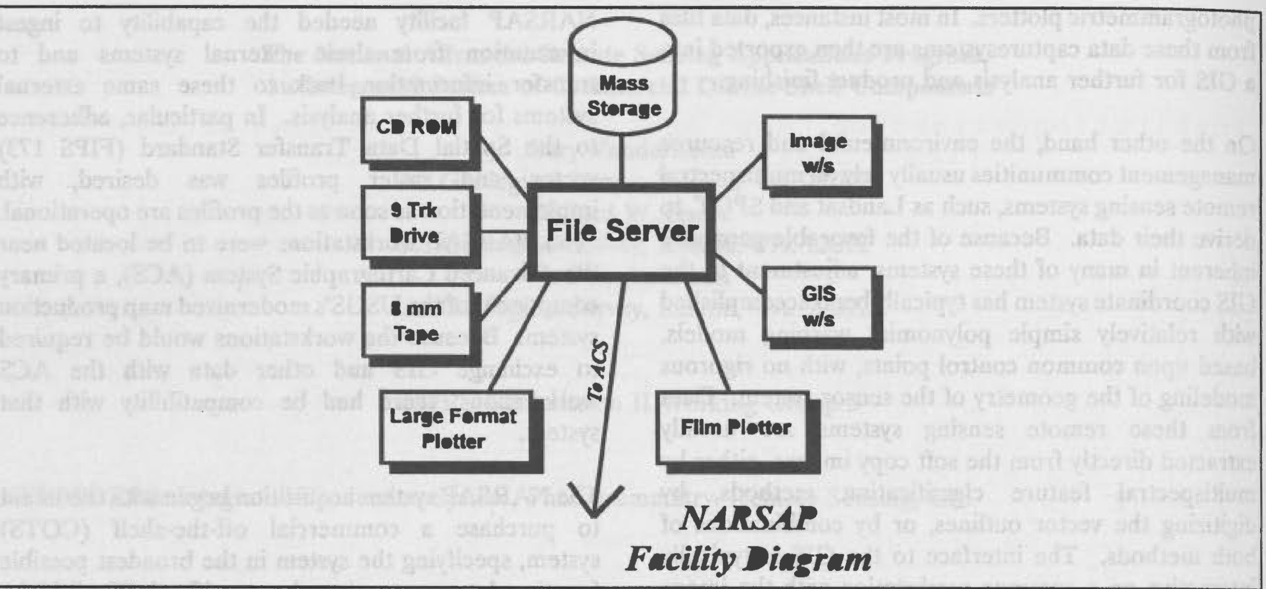


Figure 1 Proposed Configuration

Protocol Ethernet connection to other systems would be required.

Imagery to be supported would include:

- -Conventional black-and-white aerial photography
- -Color infrared aerial photography
- -Landsat MSS 1972-1979 (MSS-X data)
- -Landsat MSS 1979-present (MSS-P data)
- -Landsat TM/ETM "Fast Format"
- -Landsat MSS/TM/ETM swath data
- -Spot 1A and 1B
- -AVHRR

The imagery would be used only in digital form. Any film products would have to be digitized. It was anticipated that an appropriate scanner would be acquired as part of another modernization project in the USGS.

Other data formats that the system would have to be able to accept included:

- -Digital line graph-3 data
- -Digital line graph-O data
- -Spatial Data Transfer Standard
- -USGS digital elevation model data
- -DMA digital terrain elevation data
- -TIGER (census) data
- -Advanced Cartographic System digital orthophoto export data -Published USGS digital orthophoto data

The specifications would describe the basic processing functions in relation to image processing, mensuration and geopositioning, and multispectral classification and processing capabilities. The image processing functions would have to make efficient use of the various image types within the context and capabilities of an interactive GIS. The image processing capabilities would require exploiting all specified image types either independently or synergistically.

The system would also have to support the stereo display of images. The stereo display would be maintained through the use of rigorous photogrammetric models. Polynomial image-to-image transformations would be allowed only when knowledge of the exposure parameters was insufficient to permit derivation of a photogrammetric stereo model. Polynomials derived from rigorous photogrammetric models, though, would be used for realtime implementation of stereo models.

For photogrammetric operations, mensuration functions would include the measurement of fiducials, pass points, control points, and profiling. Digital image correlation functions would also be required for the production of digital elevation models.

Image mensuration capabilities would have to include the determination of geocoordinates and the computation of distances, areas, azimuths, volumes, and the heights of operator-selected points. Height measurement capabilities would have to include measuring heights from stereo imagery, from monoscopic relief displacement, and from the measurement of shadows.

Geometric registration capabilities would require registering an image to an image or an image to a map or performing image reprojection. All map projections included in USGS's standard coordinate transformation software, the General Cartographic Transformation Package, would have to be supported.

The image processing functions specified for the system were:

- -Image zoom
- -Image rotation
- -Image pan
- -Image compression/decompression
- -High- and low-pass filters and edge-sharpening algorithms
- -Gray-scale remapping, with both interactive and automatic default adjustments and pseudocoloring.
- -Image flicker or wipe
- -Fading of image intensity independent of the intensity of any superimposed graphics.
- -Arithmetic and logical functions, including the capability to perform basic mathematical operations on a single image or on multiple images, to combine virtual memory planes, and to create a percentage mixture of two virtual-memory planes.

The basic multispectral processing capabilities would include:

- -Manipulating bands by color plane.
- -Selecting band combinations to allow display of up to seven spectral bands and five additional derived bands all logically linked and managed as a single image.
- -Mapping any spectral band to any of the three primary colors.
- -Performing both supervised and unsupervised image classifications. Image classification functions would provide for statistical output and for postclassification smoothing and filtering of the image.

A GIS would have to be provided as part of the system. It would need to possess both raster and vector capabilities and to be fully integrated with the system's image processing capabilities. Further, it would be important for the operators to be able to use the GIS in conjunction with the displayed images without having to transfer from window to window. In other words a common user interface was desired.

The GIS would need to permit the display of vector-based GIS information, in graphics form, superimposable over both monoscopic and stereoscopic images The graphic presentation would have to be in stereo whenever the necessary elevation data were available, and when not available, it would have to be displayed at the median elevation of the model.

Although it was hoped that the system would be able to exercise the full-function set of vector GIS capabilities while viewing and manipulating an imagery model in stereo, at the very minimum the system should allow the user to select features for display by attribute, to digitize, edit, and save information in the GIS data base, and to query the GIS data base for attribute information, all while maintaining full image processing capabilities.

The GIS would also need raster processing capabilities to permit conversion between raster and vector data sets, Boolean operations within raster data sets, and display of raster data sets in conjunction with registered images, with flicker, wipe, and split-screen image processing capabilities present.

The system also must be capable of maintaining raster maps, orthophoto images, and near-orthophoto or nearly rectified images so that they could be readily displayed in answer to a geographic query.

The GIS would need the capability of incorporating new or revised functions.

The system would need enough mass storage capacity to permit storage of adequate data for analysis and to permit retrieval of that data in a timely manner. Although GIS and imagery data would not need to be integrated into the same storage scheme, they would have to be addressable in a similar manner. Finally,there would have to be an index accessible through the GIS that would index all system images, both those online and those in temporary storage.

THE NARSAP SYSTEM

The system purchased by the USGS was integrated by Autometric, Inc., of Alexandria, Virginia. All components are COTS items. The system consists of Silicon Graphics, Inc., (SGI) hardware, with ERDAS Imagine image processing software, including the Spatial Modeler, Vector Module, Image Catalog, Toolkit, OrthoMAX, and ARC/INFO GIS. The exact hardware configuration of the system as delivered consists of:

The system server, an SGI Challenge L with dual CPU's, 128 Mb of memory, a 1.2-GB system disk, 24 Gb of shared disk space, an 8mm tape drive, and a CD-ROM drive.

Five SGI Indigo Elan workstations, each with 96 Mb of memory, a 1-Gb system disk, 2-Gb of user disk space, a 19-inch monitor with Stereographics, Inc., "CrystalEyes" infrared emitter and glasses.

Also delivered as part of the system were a Tecktronix Phaser IIsd printer and a Versatec electrostatic plotter.

The server and workstations are connected through a dual-ring FDDI fiber optic network. Additionally, the server, workstations, and plotter are each individually linked to the main Ethernet network. In addition to allowing each NARSAP workstation to communicate with and transfer data to other production systems, this linkage provides backup to the FDDI should it fail for any reason. Although considerably slower in moving large image files than the FDDI, the NARSAP system at least will not be brought to a halt.

When the functional specifications were being written, one goal was that there be a single common user interface for all functionality, photogrammetric, image processing, and GIS. This goal was not fully achieved. The ARC/INFO functions were not fully integrated into a common user interface, but the link between ERDAS Imagine and ARC/INFO was made transparent by using the ARC/INFO data structure to store vector and attribution data. The vector module of Imagine will allow selective display and editing, including attribution, of vector data. The data are stored in an ARC/INFO file structure. An operator who needs to perform analysis functions on the vector data, such as polygon overlay, buffer zones, and so on must operate in the ARC/INFO HMI structure. The derived vector data sets can then be displayed in ERDAS over images.

The only requirement in the specification not fulfilled by the off-the-shelf system purchased was the capability to display vector information superimposed over a three-dimensional model. Because the GIS that was chosen does not store elevation information with each digitized point the digital elevation model is used to assign an elevation to GIS features. Features above ground level will not be displayed at their true elevation.

As the accompanying concept-of-operations diagram indicates, no simple data flow through the system takes source data and produces a product. An almost unlimited variety of ways are available to enter and move data through the NARSAP system, which was one of the goals of the design: to make as versatile a system as possible within the limitations of photogrammetric, image processing, and GIS technologies at a reasonable cost. Therefore, the concept of operation for NARSAP depends entirely on what one wants to accomplish. The output of virtually all system capabilities can be products in themselves or become intermediate steps providing data for further integration and analysis.

NARSAP APPLICATIONS

It is hoped that bringing together scientists and researchers from various disciplines will lead to the development of specifications for new remote sensors and image exploitation systems and to the better use of existing systems in providing improved capabilities for environmental assessment, for emergency response planning, and for resource monitoring and assessment.

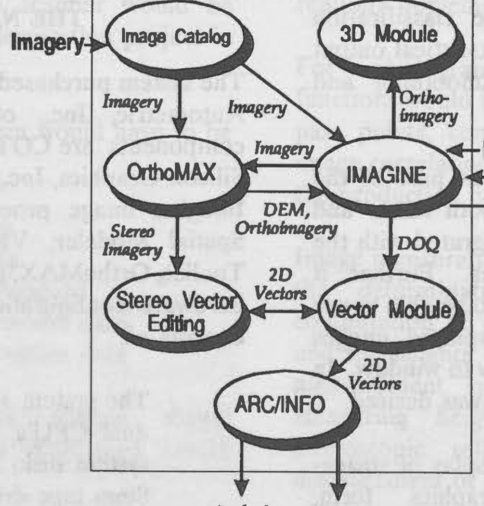


Figure 2 Operations Concept

Areas of interest identified during the system planning phase included: emergency and disaster detection, assessment, response, and prevention; climate change; environmental emergencies and protection; dam safety; crop, forest, and rangeland inventories; productivity assessments, disease and pest control; energy and minerals resource assessment; geologic mapping; wetlands inventories.

Federal civil agencies with an interest in remote sensing applications now include:

Department of Agriculture
Department of Commerce
Department of Energy
Department of the Interior
Department of Transportation
Environmental Protection Agency
Federal Emergency Management Agency
National Aeronautics and Space Administration
National Science Foundation
National Transportation Safety Board
Nuclear Regulatory Commission
Smithsonian Institution

Studies and projects proposed for the program include:

Exploring the use of remote sensing systems and analysis techniques to identify and detect possible precursors to geologic disasters such as earthquakes and volcanic eruptions.

Providing rapid assessment of a natural disaster's severity; for example, to verify personal, industrial, and governmental assistance claims. The flooding in the Midwest last spring and the Northridge California earthquake of this year are the most recent examples of the need for such assessment.

Exploring the application of remote sensing systems to detecting environmental emergencies such as oil spills and underground fires in landfills and coal mine refuse piles. The United States currently has no effective system for detecting and monitoring environmental emergencies.

Investigating the utility of remote sensing systems in inspecting dams and reservoirs for signs of potential failure or collapse. Most of the nation's dams are inspected only every 3 to 6 years. Techniques for increasing this frequency will be investigated.

Investigating techniques for the identification and inventory of crop, forest, and pasture lands along with techniques for detection and support in the control of pest infestations.

Investigating techniques for geologic mapping through obscuring forests, vegetation, and clouds as well as techniques for the automation of geologic mapping.

CONCLUSION

The concept of an integrated production system is bringing together the functions of photogrammetry, image processing, and the geographic information system in a common workstation environment. Such systems are a logical development given the continuing growth in the demand for georeferenced data by earth scientists, resource managers, and urban planners. As NARSAP demonstrates, these systems are becoming commercially available. Systems such as this one that supports the NARSAP provide users with a versatile set of tools and should promote the development of new techniques in using data from remote sensing systems. Within the U. S. Government scientists and administrators are finding them of tremendous value for investigating and managing a wide variety of environmental, resource, and emergency response management activities.

The National Advanced Remote Sensing Applications Program : un système intégré constitué de produits disponibles sur le marché

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif

DIGITAL ORTHOPHOTO
THE BASE MAP OF THE FUTURE

J.Clark Beattie

LINNET Graphics International Inc.
600-191 Broadway Avenue
Winnipeg, Manitoba, Canada R3C 3T8
Telephone: (204) 957-7566
Fax: (204) 957-7568

Abstract

As little as two years ago the concept of digital orthophotography was rarely mentioned outside of the research labs. Today, many vendors provide software to perform the corrections, soft copy photogrammetry is a reality, several companies provide full data production services and most GIS software has the capability of integrating digital orthophoto data structures.

LGI has pioneered the use of digital orthophotography as a reference data set for GIS applications being developed for the MLRIS. Digital orthophotography provides a cost effective solution for the development of GIS data bases since it can be produced in a timely manner, takes advantage of modern data collection systems and has the potential to dramatically change the way users visualize and utilize topographic data.

This paper will present the Manitoba experience form the perspective of how digital orthophotography data are being produced and used for various GIS applications. An analysis will be made of the pros and cons of these type of data as compared to conventional vector based topographic data sets. A technical analysis will also be presented to provide guidance to others who may wish to take advantage of this new technology.

L'orthophotographie numérique : la carte de base du future est arrivée

J. Clark Beattie
LINNET Graphics International Inc.
600 - 191 Broadway, Winnipeg, Manitoba R3C 3T8
Téléphone : (204) 957-7566; Facsimilé : (204) 957-7568

RÉSUMÉ

Il y a deux ans, le concept de l'orthophoto numérique était rarement mentionné en dehors des laboratoires de recherche. Maintenant, plusieurs fournisseurs offrent des logiciels pour effectuer des corrections, et les photos numériques sont une réalité. Plusieurs compagnies offrent des services complets de production de ce genre de données, et la plupart des logiciels SIRS ont la capacité d'intégrer les structures des données numériques des orthophotos.

LINNET Graphics International Inc. a été l'un des pionniers à promouvoir l'utilisation de la photogrammétrie numérique comme un ensemble de données de référence spatiale pour les applications SIG dans le cadre du programme SIRS du Manitoba. L'orthophotographie numérique fournit une solution rentable pour le développement de bases des données SIRS étant donné qu'elle peut être produite dans une courte période de temps, elle profite des systèmes de collecte de données modernes et possède le potentiel de changer de façon dramatique la visualisation et l'utilisation des données topographiques par les utilisateurs.

Cet article présentera l'expérience Manitobaine et décrira comment les données de l'orthophoto numérique sont produites et employées par plusieurs applications SIG. Une analyse des pros et contres de ce type de données sera faite dans le but de comparer des ensembles de données conventionnelles topographiques à base vectorielle. Finalement, une analyse technique sera présentée afin de guider les professionnels qui voudraient profiter de cette nouvelle technologie.

INTRODUCTION

New computer technology, improved software systems and overall computer awareness of individuals has greatly impacted the way in which we carry out daily lives. GIS technology has been the benefactor of these changes since one of the major constraints of early systems was the large amount of computer power required to make them work. Hardware and software advances have brought GIS to the desktop and the corresponding benefits to the organization are beginning to be realized.

Another major constraint to developing successful systems is providing current, accurate, reliable spatial data to serve as the foundation for GIS applications. Some vector data sets are available over a wide area, however, they usually require a substantial amount of effort to bring them into a GIS usable form. In the case of digital topographic map data the problem is further exacerbated by the fact that data which are available are often at small scales and are not up to date and therefore may not meet the spatial accuracy or content requirements of the user.

One new technology which has proven itself over the past few years is Digital Orthophoto Imagery (DOI). Data products made using DOI technology can provide a cost effective solution to the mapping problem while at the same time delivering a data set which provides substantial extra value to the GIS user.

This paper will explore the concepts behind DOI, examine the pros and cons for DOI as a GIS data set and relate the experiences to date of Linnet Graphics International (LGI) in producing DOI data for the Manitoba Land Related Information System (MLRIS).

WHAT IS DOI

DOI technology evolved out of the remote sensing community in response to the requirement to spatially correct digital data obtained from satellite sensors. In order to use these data for analysis it is necessary to remove the distortions caused by sensor operation and to provide an orthogonal picture which can be oriented to fit the earth based geodetic reference system. With these corrections performed it is possible to accurately locate the digital data relative to it's corresponding place on the earth surface and thereby enhance the useability of the data for analysis purposes. The restitution of data obtained from space based platforms requires not only the modelling of complex geometry, digital elevation terrain data and sensor performance characteristics, but also requires the manipulation of large raster data sets. These same principles, when applied to imagery obtained from aerial camera system, have resulted in the development of DOI technology. Early systems developed for this purpose required extensive computing power, however, the latest systems are taking advantage of PC processing power and are providing the rectification functionality on the desk top. Full production systems still require strong computer systems but the functionality of DOI is being found in most GIS vendor software today.

In the traditional mapping community the procedures and technology for producing hard copy orthophoto maps has been well known and used for almost thirty years. Although these products provided users with an enhanced presentation of traditional topographic features, the fact that the medium used for presentation required a photographic process limited it's real use within most organizations. The concept of having these type of presentations available in digital form was always outside of reality until software became available to analytically resolve the correction process.

In simple terms a DOI consists of a digital representation of an aerial photograph, corrected to account for the distortions in the original image due to camera calibrations, aircraft movements and terrain elevation. By removing these distortions, (also the objective of traditional photogrammetric mapping systems) it is possible to provide a digital image which is true to scale in all directions and which can be used in every way like a map.

THE DOI PRODUCTION PROCESS

The process required to produce a DOI consists of four main components, the source imagery, aerial triangulation, the digital elevation model and the rectification.

Source Imagery

Usually the source imagery consists of aerial photography taken with conventional camera systems, although any camera for which the calibration characteristics are known could be used. For best results the camera system should be equipped with forward motion compensation to minimize the affect of image motion (the movement of the plane during the instant of exposure). Also exposure conditions should be optimized to enhance the photographic quality of the imagery. Obviously the imagery must be free from imperfections, clouds and other marks which might lessen the visual quality of the resulting data set. The scale of the source imagery must commensurate with the intended result and this should be decided prior to the acquisition of the photography if possible. Traditional enlargement relationships between source image scale and final map scale (usually 4 or 5 times) do not generally apply to DOI because the whole rectification process is digital and therefore not subject to the same sources of error as found in traditional photogrammetric procedures. Accordingly, the final product must be defined in terms of pixel resolution rather than map scale. This relationship will determine how the source imagery is to be scanned at the initial stages of production. See Figure 1 for a comparison of scanning resolutions and resultant pixel dimensions at map and ground scale. With a high resolution graphics board on a PC it is possible to obtain 10 times enlargement of 2 m pixel resolution data.

Aerial Triangulation

Aerial triangulation is a process used in conventional photogrammetry to develop the relationship between the measurements made in the aerial photography and the ground coordinate system. Aerial triangulation is a well known process and is used in virtually all photogrammetric projects and provides a rigorous reference framework for DOI as well.

A new development which is beginning to become accepted as an addition to the aerial triangulation process is the integration airborne GPS with the camera system. This provides, after post processing, perspective centre coordinates in the ground coordinate system with sufficient accuracy to be used as control in the aerial triangulation and adjustment process. Ultimately on board sensors will also provide airplane altitude data and thereby enhance the mathematical solution.

The final result of the aerial triangulation process provides input for measuring the digital elevation model and for performing the geometric corrections in the DOI rectification process.

Digital Elevation Model

The digital elevation model provides the data required by the rectification software to resolve distortions in the source imagery due to elevation changes in the terrain. There are several potential sources for these data but the most common approach is to measure the elevation data in a conventional photogrammetric instrument and record these data in the format required by the rectification software. Other techniques for creating elevation data include using digital contour data taken from existing map files and auto correlation of digital imagery.

Whichever technique is used it is important that the accuracy and coverage of the terrain data commensurate with the desired end product. Major elevation differences will cause image distortions and potential scale discontinuities between corresponding areas in the image. Careful attention to the specification for collection of elevation data will ensure that the optimum result is achieved.

Rectification

The first step in the rectification process is to scan the source imagery. This provides the uncorrected digital image data in raster format at the chosen resolution. From this point on the rectification process is carried out in the computer under software control. The first step is to measure the fiducial marks which are located in the corners of each aerial photograph. These are reference marks which define the camera frame and are used to relate the camera calibration parameters to the numerical model required in the software. The second step is to measure the location of control points identified during the aerial triangulation process. This provides the relationship between the image coordinate system and the ground control system. The third step is to link the control point file to the digital elevation file since both of these files are required for input to the rectification program. The rectification program essentially performs a space resection for each pixel in the image and then applies a correction to each pixel to remove the distortions inherent in the source data. Each pixel is processed in turn until the complete image is proceeded. The resulting file is then stored for delivery to the client system. The accuracy of the solution depends on the assumptions made in the algorithms and the quality of the input data stream. Generally the accuracy of the resulting DOI, as compared to conventional mapping techniques, will be consistently higher. The final DOI product can be delivered in a number of standard raster file formats such as TIFF, COT etc.

WHERE DOES DOI FIT IN GIS

GIS technology provides the capability for users of these systems to analyze, measure and visualize geographic features and data related to these features. A critical element, in a successful GIS application, is the reference map to which all other data sets are spatially related. DOI not only provides this reference base but also provides a visual picture of all of the geographic features required by the user. Many GIS users are satisfied with being able to view the photographic data for their area of interest and to interpret those geographic features which are of interest to them. Other data themes can be readily associated with any of these features and if necessary the feature itself can be digitized directly from the image data. Since the DOI forms only one theme or layer in the GIS data base it can be turned on or off as required. The zoom function of the GIS provides the capability to enlarge the image in the view up to the limit of the image resolution. For example the 1:60,000 source images (2 m pixel resolution) produced in Manitoba have been reproduced by LGI as hard copy plot at 1:5,000 scale and can be viewed on a high resolution monitor at a 1:2,000 scale.

DOI provides the user with the full picture which often contains substantially more information for the user than an interpreted map and when this picture is combined with other key data sets such as land ownership, transportation network and hydrography the user is provided with a complete set of geographic framework data.

MANITOBA EXPERIENCE WITH DOI

During the feasibility and user requirements studies carried out by LGI for the Manitoba Land Related Information System (MLRIS) project in 1989 - 1991 it was recognized that users needed a new type of reference map data set in order to develop GIS applications. Firstly the traditional digital topographic programs were not funded to level which would provide provincial coverage in a short time period. This factor was limiting the ability of potential users to get on with GIS developments and was putting an extra cost burden on projects because each one had to create it's own map data. A second aspect which was contributing to inhibited GIS development was the fact that digital data which were available often did not meet the needs of the user and required substantial rework in order to integrate these data in the GIS solution.

During LGI's analysis of GIS business functions, users often sited the need to have access to current aerial photography as a means of updating existing maps, for interpreting other data layers and for providing geographic reference for other data sets. The prospect of having digital aerial photography, spatially corrected and useable directly within the GIS drew a great deal of interest in the user community and eventually became an element of the MLRIS concept.

The MLRIS concept developed by LGI includes the building of a sharable set of land related information data sets contained within a computer system designed for distributing these data to the user community. Several key graphic layers or themes were identified and these included land survey fabric to be used for graphically showing land ownership parcels, digital elevation data for topographic analysis, digital orthophoto image data, transportation networks, and hydrography networks. All of these layers are developed on the same geodetic reference system according to defined standards thereby guaranteeing the ability of users to spatially relate all data layers.

DOI was recommended by LGI as the primary reference layer with three main products; 2 m pixel data for agricultural areas, .2 m pixel data for urban areas and 10 m pixel data for resource areas. In order to provide users with a hands on view of DOI three projects were undertaken in 1992. Two separate areas were chosen because of their potential GIS application, one having an agricultural aspect and the other having a land use planning aspect. The third project provided urban aspects. In 1993 a further rural area was completed to provide reference mapping for a major water resource study.

For the rural projects, aerial photography at a scale of 1:60,000 was obtained under contract by Northwest Geomatics Ltd. of Edmonton, Alberta and combined with ground control information gathered from existing sources. For the urban project, aerial photography at a scale of 1:6,000 was obtained under contract by Airquest of Winnipeg, Manitoba. DOI production, using the I2S system, was contracted to Intera Technology of Ottawa.

In the rural projects the resultant data files cover 10 x 10 km on a UTM grid whereas in the urban areas the files cover 1 x 1 km on the UTM grid. The grid layout was chosen for ease in handling within the computer systems.

To date the 125 rural data files which have been produced represents approximately 10 % of the total agricultural area of the province and 65 urban files have been produced in three separate geographic areas. All of these data have exceeded the expectations of the original requirements both from the point of view of acceptability as well as the accuracy and cost. Some modifications have been made in the specifications for these data sets. For example, aerial photography is now obtained on a precision sheet centered basis in order to maximize the DOI production throughout. Additionally, the amount of ground control required is minimized by obtaining GPS positional data for each camera exposure station. Both of these changes have improved the spatial accuracy and have reduced the overall cost of the final product. Plans for 1994 have not been finalized but a proposal has been initiated for obtaining full coverage of the agricultural areas in a major project to be completed within one year.

DOI WILL CHANGE THE WAY WE PERCEIVE GEOGRAPHIC DATA

In the past, interpretation of geographic features from aerial photography has been performed by people with specific training in the theory and operation of sophisticated optical mechanical measuring systems. People have the best skills for this type of interpretation. Computers have great difficulty in recognizing patterns and associating these patterns with correct geographic features. Some advances have been made in computer assisted interpretation of binary raster images and multi spectral remote sensing data, however, classification of continuous tone or grey scale raster data is less than perfected. DOI provides the opportunity for any user to perform the interpretation they require for their own use. Utility companies, for example, have been able to improve the spatial accuracy of their facility records by digitizing pole locations directly from the DOI. Other users can readily locate their data sets by referencing key objects such as roads, lakes, vegetation boundaries and built up areas.

The spatial measurement process, which is a fundamental component of conventional photogrammetric systems, can be readily performed with GIS tools and linked directly with the associated data base for subsequent analysis. The GIS user does not need to rely on someone else to interpret the image for them however if there is concern about the integrity of data digitization then photogrammetric systems can be employed. Already several "soft copy" photogrammetric systems have been developed and are in use in conjunction with GIS technology. The concept behind these systems is the similar to that found in analytical photogrammetric systems and they provide this measurement capability on the same platform as the GIS. Intergraph for example have developed an image station which allows the user to see the image in stereo and to make real time measurements on the system which has the same basic characteristics as the GIS platform. This capability will be fully integrated in the GIS applications in the future.

Combining the digital imagery with a digital terrain model also provides the user with a three dimensional model of the geography they are modelling which closely resembles reality. This type of visualization can be a powerful tool not only for showing projected impacts of development projects but also for measuring in three dimensional space.

WHERE WILL DOI BE IN THE FUTURE

The future of DOI is clear, it will take a leadership position as the spatial reference data of choice for GIS applications. In the past two years DOI capability (at least viewing and basic manipulating functions) has been added to most GIS software, those that don't have the capability are developing it. There are DOI projects being performed all around the world and the USGS has a program in place to complete DOI coverage of the entire continental USA. Data compression techniques are being developed to improve the storage and retrieval functions of current systems. Satellite systems are currently being built which will provide optical high resolution (3 m pixel) data to users essentially on demand at a reasonable cost. One metre satellite technology will likely be licensed within the next two years. Digital camera systems will be developed for airborne applications thereby providing high resolution optical and multispectral data without having to produce hard copy intermediate products.

CONCLUSION

DOI is a technology which has gained rapid acceptance amongst the GIS user community and is poised to revolutionize the way which reference mapping is carried out. The first reason for this acceptance is obviously the fact that the technology does work and can deliver data in a cost effective manner. Another reason, which may be more subtle, is the fact that now users can develop their own view of the geographic space they are analysing without having to depend on someone else interpreting the features for them. In fact, when one looks at the business functions or process carried out by most GIS users, the traditional vector topographic map is used primarily as reference only. Many users do nothing but overlay their data for analysis, the map itself only provides the spatial reference.

DOI provides the same spatial reference with the added value of providing the current aerial photography image. Now GIS users can interpret their own features where necessary and use the image to enhance other data themes they are using for analysis. They have the added advantage that the image data is probably more current, more accurate and more flexible than the vector data they were using before.

DIGITAL REVISION OF NTS MAPS: A PILOT PROJECT

Anna Marie Regan and Costas Armenakis

Canada Centre for Mapping / Centre canadien de cartographie
Natural Resources Canada / Ressources naturelles Canada
615 Booth St. / 615 rue Booth, Ottawa, Ontario, K1A 0E9
tel.: (613) 943-0466, fax.: (613) 943-8959
e-mail: armenaki@emr1.emr.ca

ISPRS Commission II, WG II/III

KEY WORDS: Map revision, generalization, vector-raster integration

ABSTRACT

The National Topographic Series (NTS) map revision pilot project that was conducted at the Canada Centre for Mapping is described. This project investigated the potential of using 1:50 000 digital data and Landsat TM imagery to derive and revise the 1:250 000 NTS maps, and digitized aerial photography to revise the 1:50 000 NTS maps. The UNIX based CARIS system was used for both the derivation and the revision of the prototype data sets. For the derivation of the 1:250 000 map, data generalization techniques were applied to a 1:50 000 data set. The revision of vector data from both Landsat TM image and digitized aerial photography was performed by integrating the raster images and the vector map data. Change detection was performed visually and new vector data collected. The procedures followed and the results obtained are presented.

RESUME

Cet article décrit le projet-pilote de révision cartographique de la Série de cartes topographiques nationales qui fut exécuté au Centre canadien de cartographie. Le projet-pilote a étudié les possibilités d'utiliser des données numériques au 1:50 000 et des images Landsat TM en vue de recompiler et de réviser les cartes au 1:250 000, ainsi que les possibilités d'utiliser des photographies aériennes numériques pour réviser les cartes au 1:50 000. Le système CARIS, articulé sur UNIX, fut utilisé pour la recompilation et la révision des ensembles de données du prototype. Des techniques de généralisation des données furent appliquées à l'ensemble de données au 1:50 000 pour obtenir la recompilation de la carte au 1:250 000. La révision des données vectorielles des images Landsat TM et des photographies aériennes numériques fut accomplie en intégrant les images raster et les données cartographiques vectorielles. L'identification de changements fut accomplie visuellement et de nouvelles données vectorielles furent recueillies. Les procédures suivies et les résultats obtenus sont présentés.

1.0 INTRODUCTION

The Canada Centre for Mapping, (CCM), is responsible for the production and maintenance of the National Topographic Series (NTS) maps for Canada. The NTS consist of maps at two scales: 1:250 000 and 1:50 000. Canada has complete coverage with the 1:250 000 maps, which have also been scanned for conversion to digital form. Although these 914 maps are in digital form, the data is currently only being used for GIS applications, not for map revision. These maps continue to be manually revised. The coverage of Canada at the 1:50 000 scale is approximately 92% complete, with a combination of digital and analogue maps. Of these 11, 931 maps at the 1:50 000 scale, 20% have corresponding digital data, either from stereocompilation or scanned reprographic material. This percentage of digital data is increasing as more new maps are stereocompiled and existing anologue maps revised by digital recompilation or scanned and converted to digital data. All of the digital data for the NTS maps is archived in the National Topographic Database. The only digital revision currently being used is the stereo recompilation of entire map sheets. At the present time CCM is in a transition period from combined analogue and digital production scenarios for the production and maintenance of the NTS maps to a fully digital solution. With the increase in the supply of new digital data and the growing client demand for digital products, the implementation of digital map revision technology has been accelerated. This paper describes the pilot project that was carried out to investigated the potential of using 1:50 000 digital data and LANDSAT TM imagery, and digitized aerial photography to derive/revise the data required to produce the 1:250 000 and 1:50 000 NTS maps, respectively.

2.0 SYSTEM REQUIREMENTS AND SELECTION

Map revision operations from aerial photographs and satellite images require several steps: interpretation, change detection, collection of new data, and integration of old and new data in the database. One of the effective methods for extraction of new data is by the superimposition of the existing digital map data over recent aerial photographs/satellite images and update the database by collecting the new information. The superimposition and the data collection can be performed in several ways (e.g., Welch, 1989; Derenyi and Shih, 1991; Bouloukos et al. 1992). Image processing functions are also required for the enhancement of the raster images to facilitate the interpretation and change detection. The revision of the 1:250 000 maps was based on a twofolded approach. First, recent larger scale maps such as the 1:50 000 digital maps are generalized to derive the smaller scale maps. Second, satellite imagery such as the Landsat TM is used as backdrop to the 1:250 000 digital data and the detected changes are screen-digitized thus updating directly the data set. The approach tested for updating the 1:50 000 maps is digital mono-orthocompilation, that is the on-screen digitization of planimetric details from registered digital orthophotographs /mosaics. The digital orthophotographs are displayed as backdrop to the 1:50 000 digital data and the newly digitized data updates the vector data set.

These requirements are met by having a vector-based topographic mapping and GIS system integrated with raster data handling capability. The integration should extend to the operational and functional levels. The system used is CARIS (Computer Aided Resource Information System), a comprehensive GIS system produced by Universal Systems Ltd. of Fredericton, New Brunswick, Canada. CARIS was originally a vector-based GIS system, which has recently extended its capabilities to fully integrate raster type data (Derenyi, 1991; Halim et al., 1992). For topographic digital mapping and GIS applications CARIS software functionality includes: 2 and 3-D data capture, interactive editing, generalization and cartographic editing, topological creation, interactive data analysis, off-line data analysis, digital terrain modeling, semi-automated map input and application support.

The raster image requirements are addressed in the areas of image display and image analysis. To display an image the following commands are available (USL, 1992): set up the colour map, enable the drawing of raster image, select the layers to be display, draw the selected layers, disable drawing of raster images and display pixel value. Image interpretation is supported interactively by contrast enhancement using either histogram equalization or user specified histogram, piecewise linear stretch, neighbourhood averaging, and bi-directional gradient. Also, available are: image registration and rectification including resampling, ortho-image generation and image classification modules.

The system configuration for the pilot project consists of CARIS UNIX version 4.2.4 running on Sun SPARCstations 2 and 10.

3.0 REVISION OF 1:250 000 NTS MAPS

The objectives of this part of the project were to update an area of a 1:250 000 digital data set, first by generalizing the 1:50 000 data set, and secondly by revising the data set using the Landsat TM image. The area of the 1:250 000 sheet, 31I Trois-Rivières, corresponding to the 1:50 000 map 31I3 of Sorel, was chosen as the pilot site for the investigation. This area also has 1:50 000 stereo compiled data coverage. Sheet 31I is currently being manually revised in the Map Revision Section using hard copy 1:50 000 maps and Landsat TM imagery (Turner and Stafford, 1987). This sheet was chosen because of the availability of a current 1:50 000 digital file and that the results of the pilot project could be compared with the manually revised product. The materials required for the pilot project revision were the 1:250 000 digital file, 1:50 000 digital data, reference points, digital Landsat TM imagery, aerial photography and the NTS maps at both 1:50 000 and 1:250 000 scales.

3.1 Derivation through the Generalization of the 1:50 000 Digital Data

For this project the vector data files were loaded in SIF format and converted into the CARIS data files required for processing and display. Based on the review of the procedures currently being used for the manual derivation of a 1:250 000 NTS map from 1:50 000 maps, the methodology of deriving one layer at a time was used to digitally derive the pilot area. The following sections describe the steps of the derivation of each layer.

3.1.1 Review of the 1:250 000 Specifications

To determine the required content and minimum size for the derived data set, the 1:250 000 Polychrome Mapping Specifications were reviewed. Content refers to the features that are shown on the 1:250 000 map and minimum size refers to the smallest area or line to be shown. To complement the specifications which are currently being revised, the project team relied quite strongly on the directives and minimum size guide being used in the Map Revision Section.

3.1.2 Data Generalization

The generalization process used for the derivation of the 1:250 000 data set from 1:50 000 data consisted of the following eight techniques (Mackaness, 1991):

a) Feature correlation/selection:
Having identified the required content, the two data sets were first correlated in order to match corresponding features. A direct correlation of feature codes was not possible because the feature codes for the 1:250 000 data are not all the same as those used for the 1:50 000 data. Next, the content required for the 1:250 000 data set was extracted from the 1:50 000 data set Finally, the feature codes of the selected data were changed where necessary. Caution must be taken during this stage not to delete any features that are required for data continuity, such as dugouts that are located on a river. The correlated data was separated into the layers; hydrography, transportation, vegetation and wetland, culture, built-up area, contours, and gravel pits, and layer numbers assigned.

b) Data filtering:
The Douglas-Peucker filter algorithm was applied to the transportation layer at 50m and to the hydrography layer at 25m intervals. These filtering intervals were tested and chosen to ensure that the data retained its shape while reducing the amount of data points, which in turn accelerated the data handling and processing.

c) Omission of features not meeting minimum size requirement:
Using the minimum size guide, features that were too small, too narrow, or too close together were identified and interactively removed from the data set. The exception being those features that could be grouped together or that were required for data continuity such as transformer stations on power lines.

d) Grouping of features:
Polygons not meeting the minimum area requirement were interactively grouped together or buffers were generated surrounding them.

e) Collapse of features:
The collapsing technique was first used when a feature did not meet the minimum size requirement to be shown as one type (i.e., an area) but could be represented as another type (i.e., a point). In the cases of a double line river that did not meet the minimum size, a new line was digitized representing the river center line, and a campground area that was not large enough to satisfy the minimum requirement was deleted and replaced with a point type. A second situation where collapsing was required was where a feature, such as an interchange was represented as one type at the 1:50 000 scale, but as another type at the 1:250 000 scale.

f) Combination (reclassification) of features:
Reclassifying features for generalization means changing their feature code in order to connect the feature with an adjacent feature. Segments of lines not meeting minimum length requirements such as ditches at the end of a stream were reclassified to stream.

g) Simplification (line smoothing filter) of data:
Line smoothing was applied to features that were jagged to make them cartographically presentable. It was necessary to smooth the streams from the 1:50 000 data for presentation at the 1:250 000 scale.

h) Displacement of features:
When decreasing the presentation scale of data, features often conflict with each other and must be displaced for cartographic representation. An example of a feature being displaced in this data set was a rock symbol that had to be interactively moved because it was too close to the shoreline at the 1:250 000 scale.

3.1.3 Building Topology

The requirement for clean structured data for this type of an application was apparent. Since the data set was not cleaned or structured, it was decided that sample cleaning would be done in order to utilize some of the automated generalization functionality. The two features that were selected for the sample cleaning were roads and gravel pits. The following is a brief description of the cleaning performed and the generalization applied to the clean data:

- The roads were cleaned by building their topological network, which deleted overshoots, snapped undershoots, and removed pseudo nodes and duplicated lines. This allowed the file's arcs (roads) to be categorized by length, the different categories to be highlighted on the display, and the deletion of a category of lines interactively or in batch.

- Since the gravel pit file was relatively small, its full polygon topology was built. This enabled: all the polygons below minimum size to be highlighted on the display, the grouping of the ones that belonged together, and the use of the "dissolve polygon" CARIS functionality to automatically delete the remaining polygons smaller than minimum size.

3.1.4 Edge Matching of Map Data Sets

Edge matching is a term used to describe the process of making two adjacent data sets graphically continuous so that the features on both sheets continue smoothly from one data set into the other without breaks or disjoints. This functionality would be required for joining the sixteen 1:50 000 digital files that make up a 1:250 000 data set to create a seamless file. There were many options available in the CARIS edge joining routine and, depending on the application, the operator can customize the routine with the different option qualifiers. Some of the options were as follows: edge matching within one file or between two files; matching lines within a specified maximum distance (map or ground distance); restricting the processing to a given layer number(s) or feature code(s); and joining lines with the same contour values. Also, the operator can specify a method of joining the lines such as straight extend (straight line between end points) or modify the lines so the join is made smoothly. The results of the execution of the routine were displayed and found that approximately 60% of the lines had joined. Since the tolerance must be kept low to avoid incorrect joins, there will always be the need for interactive editing.

3.1.5 Metrication and Editing of Contours

Since the contours from the 1:50 000 data set were metric and those from the 1:250 000 data set were imperial, the decision to use the metric contours was consistent with CCM's policy with respect to the requirement for all new products to be in metric. A 20m interval would be sufficient, thus every second contour line was selected from the 1:50 000 data set. Editing of contours was required for the re-entrance with single line rivers that had been generalized from double line rivers and to ensure that they did not enter the newly derived gravel pits.

3.1.6 Quality Control

Quality control is more a procedure that must be set, than a functionality that can be specified. The initial step in the quality control procedure was to run a statistical listing for the derived data set and to check that the feature codes were correct for the 1:250 000 product. Plots of each derived layer were then produced at the 1:250 000 scale and inspected for content, minimum size and cartographic presentation. These plots were also overlaid on each other and checked for fit. The results of the comparison of the manually revised map sheet and the derived data plots were very favorable with slight variations due to generalization choices, within the accuracy tolerances.

3.2 Revision Using Landsat TM Imagery

The 1991 Landsat TM imagery that was acquired consisted of seven spectral bands and covered approximately the area of a 1:250 000 map sheet. Landsat TM has a resolution of 30m. The image was subdivided and reformatted to PIP (PCI-BSQ intermediate file format) by Canada Centre for Remote Sensing on the PCI System and input to the CARIS System. The CARIS System then reformatted the PIP file to the CARIS IPV format which was required for processing and display.

3.2.1 Measurement and Evaluation of Reference Data

In preparation for the measurement of the reference points, spectral bands 7, 2 and 1 of the Landsat TM image were selected and the image was enhanced by performing contrast stretching through a histogram equalization on the red, green and blue bands. This enhanced the detail in the image and improved the identification of the selected control points. To collect data from a Landsat image, the image must first be registered to projection reference system to ensure that collected data is in its proper planimetric position. Due to the small mapping scale, low relief and immediate unavailability of DEM, correction for relief displacement was not performed for the TM image. The registration of the image to the vector data was performed by a polynomial type transformation (Welch et al., 1985; Colvocoresses, 1986).

Thirty-five (35) reference points were measured interactively in both the vector file and Landsat raster file. In some cases, digitizing these reference points proved to be difficult, especially where the point was located in an area of similar spectral reflectance such as a rural road intersection surrounded by cultivated fields. An interactive evaluation of the control and reference points was performed. This functionality allowed two data sets to be graphically displayed and their respective control files to be evaluated during the adjustment. When evaluating the control points, the operator may remove points with large residuals, add new points, re-run the same adjustment or perform the adjustment using another transformation. The geometric correction of the Landsat TM image was performed using a third degree polynomial transformation. Twenty-one (21) control points were selected for the rectification of Landsat image and the registration to the vector data. The standard deviation of the planimetric residuals after the adjustment was ±14.7m (±10.9m in x and ±17.7m in y). Nearest neighbour resampling was applied to the image and the output pixel size was kept at 30m. To obtain an indication of the quality of the adjustment, the coordinates of fourteen (14) image check points were measured compared to their "true" ground values. The standard deviation of the differences between "true" and measured coordinates were calculated and for ten of the check points were ±12.6m in x and ±16.1m in y. The coordinate differences in the other four check points varied between sub-pixel values to two-pixel values. These results satisfy the NATO A rating planimetric accuracy (CMAS=125m) for the 1:250 000 maps, equivalent to standard deviation of differences of ±41.2m.

3.2.2 Change Detection and Vector Data Collection

The task of change detection was performed visually by displaying the integrated vector and raster files to determine areas within the vector file that have experienced physical change. Only minor changes, such as the enlargement of a couple of gravel pits and the addition of a road were detected. To provide an opportunity to try the revision functionality, features from the vector file were masked to simulate change.

For the collection of the 'updates', a unique layer number was set to differentiate the data collected on the Landsat TM image from the original vector data. This was necessary for later separation of the Landsat data for comparison to the original data. Data collection was performed interactively using screen digitization (mouse & cursor). The features collected were: roads, built up areas, water courses, waterbodies, gravel pits, forests and power lines. The identification of features was considered biased, as their identity was already known from the masked data. It was found that in areas of high contrast the feature delineation was highly accurate. Examples of features that were relatively easy to interpret were power transmission lines cut through forested areas, water body boundaries (but not narrow streams), most forest areas, gravel pits and multi-lane roads. Classification of roads and the differentiation between cut lines, roads and power lines would not be possible directly from a Landsat TM image and would require field verification or the use of ancillary data. Cases where it was difficult to determine the feature delineation were between rural roads and cultivated fields, built up area extent, residential roads through built up areas, and between some forests and fields. Since digital aerial photography was available, it was referenced to clarify the ambiguous areas. A comparison of the data collected from the false color image using bands 4, 3 and 2, to data collected from the true color image using bands 7, 2 and 1, showed that the operators obtained better results from the true color image.

3.2.3 Quality Control

Plots of the original derived file (before the features were masked) and of the data collected from Landsat were produced at the 1:50 000 scale. Since both the derived data and the data collected from the Landsat image were registered using the same control, the comparison to check their geometric fit was made visually by overlaying the plots. It was found that the collected Landsat data was jagged due to the point to point collection method used. The Douglas-Peucker filter and a smoothing operation were applied interactively to the Landsat data, which brought the two files to a closer geometric fit. The maximum spatial discrepancy between the two data sets existed in gravel pits and forest areas. These are two types of land uses that have the potential to need frequent updating and are not likely errors but valid changes. Linear features and water boundaries were found to deviate from their position in the derived file often by less than 50 meters. When the plots were done at the 1:250 000 scale the linear feature discrepancies are almost negligible. It was found that the collected Landsat data matched the derived data within the accuracy and tolerance required for the 1:250 000 map.

3.2.4 Production of Paper Maps

The capabilities of CARIS for the integrated vector/raster data output for the production of image maps were also evaluated. Cartographic editing, placing toponymic data and contour labels, generating the border, UTM grid and map surround, for the production of line maps is performed by the CARIS cartographic editing system. The cartographic vector data was exported to POSTSCRIPT format for input to Products and Services Division's, (PSD), SCITEX system. The co-registered raster data was exported in PIP (PCI-BSQ intermediate file format) for input to PSD's PCI system, where it was prepared for plotting. To achieve good visually accepted image print, the TM image was resampled using cubic convolution with output pixel size of 25m. Image maps at 1:50 000 and 1:100 000 scales were produced on the IRIS plotter and were composed of the Landsat TM image, contours from the vector file, and the border and grid.

4.0 REVISION OF 1:50 000 NTS MAPS

An experiment was also performed to evaluate the potential of updating the 1:50 000 topographic maps from digitized photography. Both digital rectified and orthorectified photographs were used for the test. The 1:50 000 data set had been stereo compiled from 1984 photography. This data set contained unstructured positional data as it was compiled, not the cartographically edited data as presented on the map. The vector data files in SIF format were loaded and converted into the CARIS data files required for processing and display. Aerial photography at scale of 1:50 000 was used to evaluate the accuracy of the digital revision process. The Helava DPW770 system was used for the scanning and the production of digital orthophotos. The diapositives were scanned at 25 microns (1000 dpi). A 50m DEM grid was generated and used to create the orthophotos. The image was resampled with nearest neighbour with an output pixel size of 2m ground resolution (about 40microns). Both digitized photography and othophotography were exported in DOS TIFF format and using the CARIS data exchange utilities were converted to CARIS IPV raster format.

4.1 Measurement and Evaluation of Collected Data

The number and location of the control points were based upon the points used for the aerotriangulation, which covered the extent of the digital photography. For registering the orthophotograph to the "old" map digital data, an affine transformation was used. The digitized photograph was rectified using the projective transformation. A total of nine (9) control points were used for the transformations. With the raster and vector images combined, the coordinates of thirty three (33) check points were measured in the raster image. These points were evenly distributed throughout the registered area and were well identifiable. Their ground coordinates were compared with the "true" values of the existing positional data. For the rectified photograph the standard deviation of the differences was ±12.5m in x and ±16.0m in y. For the orthophoto the standard deviation of the differences were ±2.9m in x and ±4.9m in y.

The results from the orthophoto satisfy the NATO A rating planimetric accuracy (CMAS=25m) for the 1:50 000 maps, equivalent to standard deviation of differences of ±8.3m.

Although no extensive tests were performed at these stage, the overlay between vector data and the digital orthophotography demonstrated a good fit between the two types of data. The visual identification of changes was much easier than the one from the satellite image and was supported very well with the image processing capabilities of the CARIS system.

5.0 CONCLUDING REMARKS

As a result of the map revision pilot project, CCM's in-house expertise in vector data generalization, raster image processing, and raster and vector data integration has increased. Having completed the derivation of the pilot area through the generalization of a 1:50 000 data set, the requirements for data generalization to produce 1:250 000 data are very clear. In order to pursue this methodology, sixteen 1:50 000 digital data sets covering the 1:250 000 map are required. Although the pilot site was revised using Landsat TM imagery and the accuracy requirements are met, the team has some reservations concerning the sole use of Landsat TM imagery for 1:250 000 map revision. These reservations are based on the uncertainty in identifying and collecting certain features. Ancillary information will be required to improve the feature identification process (aerial photography, existing maps, field checks, etc.), as well as the use in the future of image segmentation techniques. The combination of the Helava system for the production of digital orthophotographs with the CARIS system offering real vector/raster data integration for the revision of existing digital spatial databases proved to be quite successful. It can be effectively implemented to a full production and quality assurance system. The production of image maps is also possible. A more thorough evaluation of this system integration and application for map revision from digital orthophotographs is in progress with the revision of an entire 1:50 000 map sheet.

ACKNOWLEDGMENTS

The contributions of the team members Gilles Auger, Art Dow, Rolland Dussault, Benoit Frappier, Hélène Lachance and Jeffrey Young are gratefully acknowledged.

REFERENCES

Bouloucos T., R. Kunarak, K. Tempfli (1992) "Low-cost feature extraction from aerial photographs for database revision", Inter. Archives of Photogrammetry and Remote Sensing, Vol. XXIX, Part B4, Com. IV, pp. 493-498.

Colvocoresses A.P. (1986) "Image mapping with the Thematic mapper", Photogrammetric Engineering and Remote Sensing, Vol. 52, No. 9, pp. 1499-1505.

Derenyi E.E. (1991) "Design and development of a heterogeneous GIS", CISM Journal ACSGC, Vol. 45, No. 4, pp. 561-567.

Derenyi, E.E, T.Y. Shih (1991) "Digital map revision in a vector/raster environment", internal paper, University of New Brunswick.

Halim M., E. Derenyi, R. Pollock, T.Y. Shih, C.K. Xian (1992) "Design and implementation of a GIS based digital image processing system", Inter. Archives of Photogrammetry and Remote Sensing, Vol. XXIX, Part B4, Com. IV, pp. 776-780.

Mackaness W.A. (1991) "Integration and evaluation of map generalization", in Map Generalization edited by B.P. Buttenfield and R.B. McMaster, pp. 217-226.

Turner A.M., D.R. Stafford (1987) "Operational revision of national topographic maps in Canada using Landsat images", ITC Journal 1987-2, pp. 123-128.

USL (1992) "Guide to using image data in CARIS", Universal Systems Ltd.

Welch R., T.R. Jordan, M. Ehlers (1985) "Comparative evaluation of geodetic accuracy and cartographic potential of Landsat-4 and Landsat-5 Thematic mapper image data", Photogrammetric Engineering and Remote Sensing, Vol. 51, No. 11, pp. 1799-1812.

Welch R. (1989) "Desktop mapping with personal computers", Photogrammetric Engineering and Remote Sensing, Vol. 55, No. 11, pp. 1651-1662.

GEOLOGICAL MAP PRODUCTION USING GIS SOFTWARE

Gary Labelle, Paul Huppé,
Vic Dohar, and Mario Méthot
Geoscience Information and Communications Division
Geological Survey of Canada, Natural Resources Canada
601 Booth Street, Ottawa, Ontario, K1A 0E8

KEY WORDS: GIS, Cartography, Geology, Topography, Scanning, Digitizing, Hardware, Software

ABSTRACT

Our primary goal is to produce geological maps using digital outlines of geology from AutoCad, FieldLog, scanned images, and other systems, incorporate digital base map information with geology, add text, line work, symbols, surround information, and output the results on colour electrostatic plotters and film imaging systems. Work has included the development of routines to handle external digital sources of base data, development of archiving procedures, cartographic editing routines, CMYK colour modelling specifications, font handling procedures, and plotting procedures to obtain hard copy output from postscript files.

CARTOGRAPHY SECTION

The section's map production procedures are based on the Environmental Systems Research Institute's (ESRI) ARC/INFO 6.1.1 computer assisted mapping and geographic information management system. The primary platform used is the SUN SPARC2 work station. Work stations are linked in a Local Area Network (LAN). This configuration offers the optimum combination of digital mapping applications, workstation performance, and network environment.

ARC/INFO software links cartographic data in vector files (ARC) with attribute information (INFO). ARC and INFO are both independent and fully integrated. Data files can be separately manipulated, and updating one system automatically updates appropriate data in the other system.

Routines and expertise that will greatly facilitate future map production is continually being developed in cooperation with the Computer Technology Section. These routines, menus or graphical user interfaces (GUI), include procedures for creating a watermask, labelled neatline, key maps, legends, scale bars, and streamlining routine tasks such as attributing geological polygons and adding complex structural symbols by providing AML routines to prompt and present logical choices. Symbol libraries for various fault types, colour tables for shading polygons, and geological symbols are being created or imported from other systems and stored in our symbol library.

THE COMPUTER TECHNOLOGY SECTION

The Computer Technology Section provides expertise by testing and evaluating hard copy output procedures, developing aids to import external digital data sources, and advice on the acquisition and future direction for automation.

WORKING WITH DIGITAL TOPOGRAPHIC BASES

Users should acquire digital base imagery from a recognized cartographic mapping institution such as NRCan, provincial mapping agencies, and private sector companies with a good track record for supplying quality digital mapping imagery. If you change the data resolution by generalizing or weeding out vertices to make the base digital imagery more manageable on PC based systems, be sure to retain the original copies for Cartography to bring the imagery back up to pre-press quality. Publication submissions should include digital bases and geological outlines. We will mix traditional and digital technologies to produce hard copy maps.

PREPARING HARD COPY MANUSCRIPTS
FOR DIGITAL PRODUCTION

A. Ink all geological contacts and faults on a photolysis image or a matte film overlay using drafting pens as follows:
  - Contacts (#000 pen); Faults (#2 pen).
  - Ink density should be as black as possible, linework should be sharp, clean and unbroken.

- Cartography will attribute contacts and faults after scanning.
- Contacts in parallel should be no closer than the ink width of the contact (i.e. line weight 0.012", the distance between the parallel contacts should be no less than 0.012")
- Close geological contacts through open bodies of water; all contacts should be closed polygons.
- Tie edges to adjoining sheets.

B. Structural symbols and reference letters must be placed on a separate overlay.

C Any additional information such as radiocarbon date boxes, area showing drift cover, etc. should be on a separate overlay.

D. Compile on a photolysis or reverse cronaflex image, and include at least 4 reference points (tics) at known latitude and longitude co-ordinates for maps at both 1:250 000 and 1:50 000 scales and every 1 or 2 degrees (including the interior co-ordinates for smaller-scale mapping). Ensure that the manuscript is carefully registered to the base. Indicate each overlay by map name, theme (i.e. contacts), a date, and a number identifying the overlay as 1 of 4. Maximum size is 40" x 70".

PREPARING GEOLOGICAL MANUSCRIPTS FOR DIGITIZING USING A DIGITIZING TABLET

It is sometimes faster to use a digitizing tablet to convert imagery to digital data. This image must be carefully registered to planned publication base. How data is converted depends on what hardware and software is available, and on the expertise available to use it. At GICD Cartography, scanning technology is used as much as possible. Large format digitizing tables are available to add new or edited line work to previously scanned and vectorized imagery.

PREPARING DIGITAL GEOLOGICAL SYMBOLS FOR PLACEMENT ON THE MAP

This can be a challenging task for both the geologist and cartographer. There are hundreds of structural symbols used on geological maps, and the Cartography Section is only just beginning to build its own symbol library. Each symbol must be designed and digitized into a template, so that it becomes an object which can be positioned and rotated at a referenced point location on the map. With ARC/INFO, the symbol is designed and stored as a "marker symbol" that can be scaled and rotated. The author can supply an ASCII or DXF file containing location, type, and angle of each symbol. Cartography then adds the symbol to its library. Cartography is cooperating with other GSC divisions and also with the USGS to accumulate an extensive symbol library. This will require considerable time and effort. It has taken years to build up he library of templates for traditional geological maps, and it will require effort and resources to put all these symbols into a digital format. The approach is to create the new digital symbols on an "as required basis" for new publications and then to add them to the library of symbols. Authors' symbols can be submitted to Cartography in a digital format which is then imported into ARC/INFO.

PREPARING GEOLOGICAL MAP DATABASE DEFINITIONS

Guidelines for this file and coverage specifications for a digital geology database are being prepared by the GICD Computer Technology Section. If you wish to have a preliminary copy or some input into these guidelines, please contact either Cartography or the Computer Technology Section

DIGITAL BASE SOURCES

1. NTS 1:250 000 and 1:50 000 bases. These can be ordered from NRCan (613-995-0314).
2 ESRI, DCW, 1:1 000 000. These bases are available to GSC users for geological map production. The coverages are in geographic co-ordinates of 5 degree by 5 degree tiles.
3. National Atlas 1:7 500 000 and 1:2 000 000 bases in ARC/INFO format.
4. ESRI, Arc/World at 1:25 000 000 and 1:3 000 000 in ARC/INFO format.
5. ESRI, Arc/USA at 1:2 000 000 in ARC/INFO format.

DIGITAL GEOLOGY SOURCES

1) CAD DXF from PC AutoCad systems and MAC CAD systems. These files are imported into ARC/INFO with the DXFARC command.
2) DLG (digital line graph) from Intergraph or SPANS are imported with the DLGARC command.
3) Raster files (TIFF, RLC, SUN raster file)
4) ARC/INFO software
5) Spans software

HARDWARE AND SOFTWARE

1) One Sparc10 Server
2) 18 Sparc2 workstations
3) 19 Arc/Info 6.1.1 licenses
4) 1 PC-ARC/INFO 3.4d running on a 486.
5) 5 Island Graphic licenses
6) 4 AutoCad stations, 1 with ArcCad overlay
7) 3 Tektronix X terminals.

CONTACTS

Chief, Cartographic & Reproduction Services,
John Bill..................................613-996-9348

Head, Cartography Units
Vern Foster............................613-996-9358
Gary Labelle. 613-943-0738

Computer Technology Section
John Glynn. 613-996-0168
Paul Huppé.............................613-943-0996
Rob Burns. 613-943-0996

Production de cartes géologiques à l'aide d'un logiciel SIG

RÉSUMÉ

Traduction non disponible pour cause de livraison tardive du résumé définitif

Evaluation of Digital Elevation Modelling and Ortho-image Production from Airborne Digital Frame Camera Imagery

Alexander Chichagov*
Douglas King**
Department of Geography
Carleton University
1125 Colonel By Drive
Ottawa, Ontario
K1S 5B6
Tel.: (613) 588-2600 (8439); Fax: (613) 788-4301

ABSTRACT

The objective of this research is to evaluate airborne digital frame camera imaging in digital elevation modelling and ortho-image production for areas of natural and rugged terrain at scales similar to standard mapping scales. A Kodak Megaplus 1.4 digital frame camera with 1280 x 1024 pixel format was flown over the Gatineau Park in May, 1993. The altitude selected was approximately 3,940 m above ground level to provide similar ground coverage to standard 1:10,000, 23 cm photography. The corresponding ground pixel size was about 2.0 m. For an initial stereo triplet, over fifty points were selected in the imagery, four for control and the rest as check points. These were accurately surveyed using rapid-static differential GPS techniques. A digital elevation model (DEM) for the triplet was created from the control points using commercial software which includes automated image matching and elevation determination through calculation of parallax at each point. Residuals for each check point in relation to the elevation model and an overall RMS error were computed. An ortho-image for the area was then produced from the DEM. Methodology and results of airborne imagery acquisition, DEM and ortho-image production, and DEM accuracy evaluation will be presented.

Notes: Poster presentation preferred in GIS symposium, although both an oral presentation and ISPRS presentation would be acceptable if you need a paper to fill in a time slot for a specific theme. An abstract will be submitted in French with the final version of the paper.

---

* M.A. Candidate
** Assistant Professor

Évaluation de la modélisation altimétrique numérique et de la génération d'ortho-images par caméra numérique aéroportée

RÉSUMÉ

La présente recherche a pour objet l'évaluation des images obtenues au moyen d'appareils de prise de vues aériennes à trame numérique, utilisées pour la modélisation altimétrique numérique et la génération d'ortho-images de régions naturelles et accidentées à des échelles comparables aux échelles cartographiques types. En mai 1993, on a survolé le parc de la Gatineau à bord d'un appareil équipé d'une caméra Kodak Megaplus à trame numérique de 1,4 et de 1280 sur 1024 pixels. On a maintenu une altitude d'environ 3 940 mètres au-dessus du sol afin d'obtenir une couverture équivalant à celle d'une photographie de 23 cm à l'échelle de 1/10 000. Le format de pixel correspondant au sol était d'environ 2,0 m. Pour obtenir un triplet pour stéréo initial, on a choisi plus de cinquante points des images, dont quatre pour le canevas du terrain, les autres servant de points de contrôle. Ces points ont fait l'objet de levés de précision à l'aide des techniques de positionnement par GPS en mode différentiel rapide-statique. On a constitué un modèle numérique d'altitude (MNA) pour le triplet à partir des points de contrôle en se servant d'un logiciel commercial possédant des fonctions d'appariement d'images automatisé et permettant de déterminer les altitudes par un calcul de la parallaxe en chaque point. On a calculé pour chaque point de contrôle l'erreur résiduelle par rapport au modèle d'altitude ainsi que l'erreur quadratique moyenne globale. On a ensuite généré une ortho-image de la zone à partir du MNA; on présentera les résultats des opérations d'acquisition d'imagerie aérienne, de MNA et de génération d'ortho-images ainsi qu'une évaluation de la précision du MNA.

AN INTEGRATED PACKAGE FOR THE PROCESSING AND ANALYSIS OF SAR IMAGERY, AND THE FUSION OF RADAR AND PASSIVE MICROWAVE DATA

Bernard Armour, Jim Ehrismann, Frank Chen, Gord Bowman
Elena Berestesky, David Adams, Andrew Emmons
Atlantis Scientific Systems Group Inc.,
1827 Woodward Drive, Ottawa, Ontario,Canada, K2C 0P9
Julius Princz
Canada Centre for Remote Sensing, Special Technology Division, National Resources Canada
588 Booth Street, Ottawa Ontario, Canada, K1A 0E9
René O. Ramseier
Atmospheric Environment Service, Ice Operations and Research Group,
P.O. Box 100, 195 Westbrook Road, Stittsville, Ontario, Canada, K2S 1A2

ABSTRACT

EarthView SAR Application Package or *EV-SAR* is an integrated software package designed for the processing, analysis, and fusion of Synthetic Aperture Radar (SAR) data. The development to date has resulted in software that performs many functions with SAR imagery ranging from image quality measurement analysis to fusion of SAR imagery with other data sets. These include SAR image radiometric and geometric corrections, SAR image enhancement filters, SAR image display and measurement techniques, polarimetric SAR functions, and multisensor data fusion. New users of SAR imagery have found the package to be useful as a tutorial on understanding SAR remote sensing and as an introduction to SAR processing and data manipulation. Expert users have found EV-SAR to be a useful tool for working with SAR data.

KEY WORDS: Image Processing, Synthetic Aperture Radar, Geocoding, Data Fusion, Remote Sensing

1. INTRODUCTION

This paper describes the commands developed as part of the EarthView Synthetic Aperture Radar package or *EV-SAR*. EarthView is a general image processing package for PC 386/486/Pentium computers. It is used in many applications where variable image size, large dynamic range, and multiple data types (integer, floating point, complex) are required. EV-SAR is an application package built using existing commands in the EarthView command library plus a set of new commands functions written specifically for processing Synthetic Aperture Radar (SAR) imagery. Since the SAR sensor has a point target response which extends in two dimensions with a predictable sidelobe structure, and since SAR is susceptible to multiplicative noise as well as the usual additive noise, special processing tools taking into account these characteristics will aid interpretation. Several of these functions have been implemented in EV-SAR and are described in the following.

The paper is organized as follows. In Section 2, SAR image corrections are described, followed by SAR image enhancement techniques in Section 3. In Section 4, SAR image display and measurement functions are described. In Section 5, the polarimetric SAR functions are introduced. In Section 6, some multisensor data fusion functions are described, followed by the conclusions in Section 7.

2. SAR IMAGE CORRECTIONS

2.1 Antenna Pattern Correction

The variations in radar cross-section pattern, distance and incidence angle along a range line may cause severely uneven power distribution in the SAR image. The EV-SAR antenna pattern correction functions are designed to remove this effect.

The antenna pattern is estimated by averaging lines of constant azimuth to extract a mean range profile. The extraction of the antenna pattern can be done either by low pass filtering the profile or by fitting the profile to a polynomial curve. Both operations use standard EarthView commands. This approach is based on the fundamental assumption that the scattering properties of the area of interest are stable, or the variation is well characterized so that the extraction of the profile is of practical significance. The effectiveness of this algorithm is demonstrated in the example of Figure 1.

2.2 Geometric Transformations

When studying SAR imagery, it is in many instances necessary to transform a SAR image in slant range to ground range, and vice versa. EarthView employs a collection of functions for such geometric transformations for both airborne and spaceborne SAR imagery, assuming flat and spherical Earth models, respectively. In the airborne case, input parameters are altitude, stand off distance and range sample spacing; in the spaceborne case, input parameters are altitude, sur-

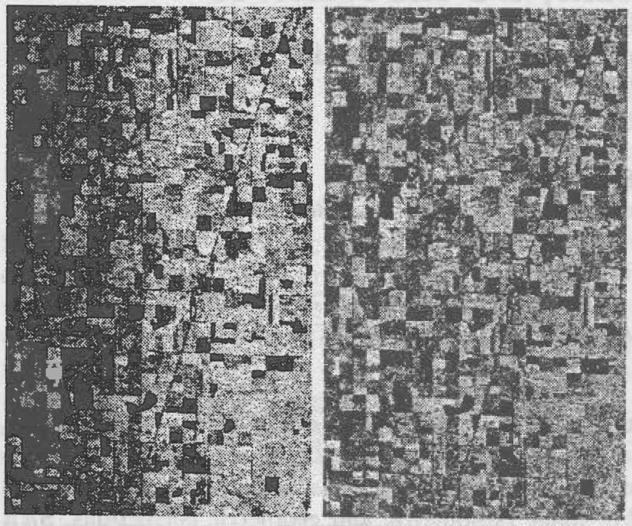


Figure 1: A SAR image before (left) and after (right) antenna pattern removal. (SAR image was produced by the CCRS C/X Airborne SAR, courtesy of CCRS, NRCan, Canada.)

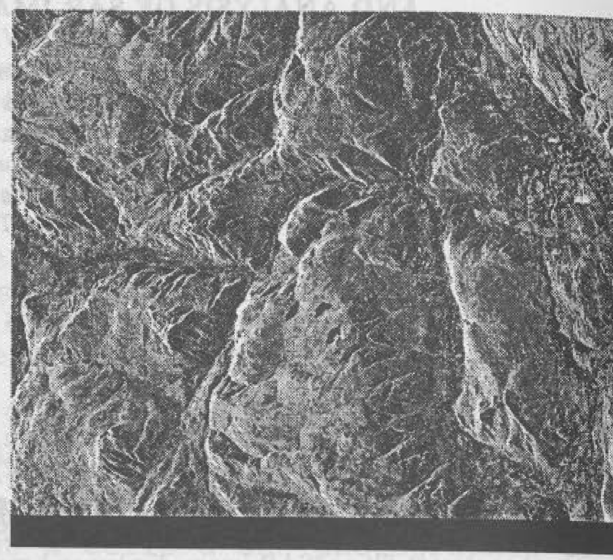


Figure 2: A section of an ERS-1 SAR image near (119E,50N). The high relief resulted in noticeable layover. (©ESA 1991)

face height relative to the average Earth radius, near range, and far range.

2.3 SAR Simulation and Geocoding

Geocoding (a SAR image) is the process of transforming a SAR image into an image whose coordinates are based on a geographical coordinate system, such as Universal Transverse Mercator (UTM) or latitude-longitude sampling grids.

SAR images of mountainous terrain are very sensitive to the terrain elevation. The range-dependent, geometric parallax found in these images makes it difficult to accurately interpret the shapes of topographical features. Significant variations in image gray level associated with geometric foreshortening confuse and in some instances completely mask image contrasts associated with changing surface cover. These phenomena make it difficult to understand and interpret such scenes, especially for interpreters trained in optical image analysis.

To correct for the induced terrain based image distortions, SAR imagery of irregular terrain elevation is *geocoded*.

The EV-SAR geocoding algorithm was adapted from the geocoding algorithm and software developed by Bert Guindon of the Canada Center for Remote Sensing [2]. It is based on prior knowledge of the terrain (in the form of Digital Terrain Elevation Data (DTED) or Digital Elevation Model (DEM)) and the parameters of the SAR platform. In the process of geocoding, a simulated SAR image is generated. This image is then co-registered with the real SAR image of the same terrain. Using this mapping and the terrain-to-SAR relationship already established, a geometric transformation is calculated, and the real SAR image is remapped, which produces the elevation-effect corrected image.

Optionally, other useful auxiliary data such as shadow and layover masks can also be generated (these characterize the interaction between the SAR viewing geometry and the local terrain orientation).

Figures 2 through 5 show some of the data and result images in an example geocoding session.

3. SAR IMAGE ENHANCEMENT

Coherent processing of SAR data makes images sensitive to small range variations on the order of the radar wavelength. These variations appear as speckle noise (manifested as an apparently random distribution of conspicuously bright or dark pixels) in the SAR image. The presence of speckle in an image reduces the ability to resolve fine details. It also degrades classification accuracy. EarthView provides a number of tools for the removal of speckle noise, as well as other types of noise. Images can be further enhanced using EarthView through histogram modification as well as classification.

3.1 Speckle Filters: Frost, Lee, MAP, MAP-Refined.

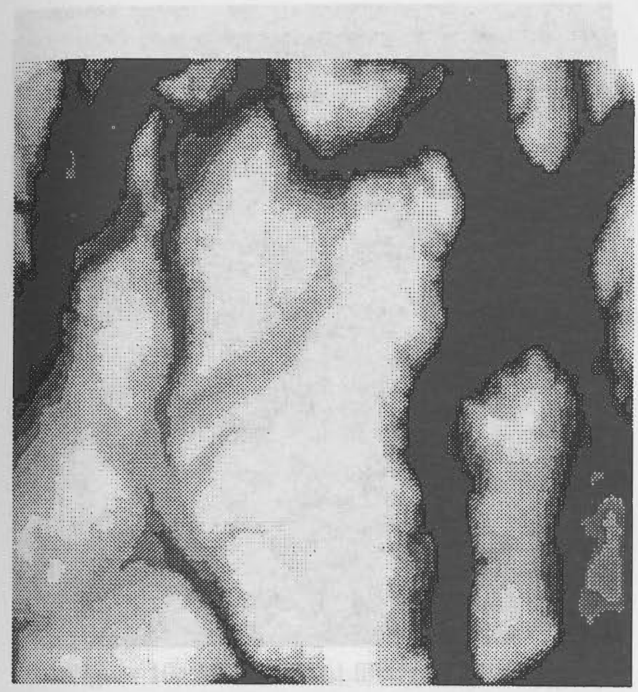


Figure 3: A subregion of the DTED corresponds to the SAR image shown in Fig. 2. Data is sampled in a lat-long grid with spacing of 3' and 6', respectively.

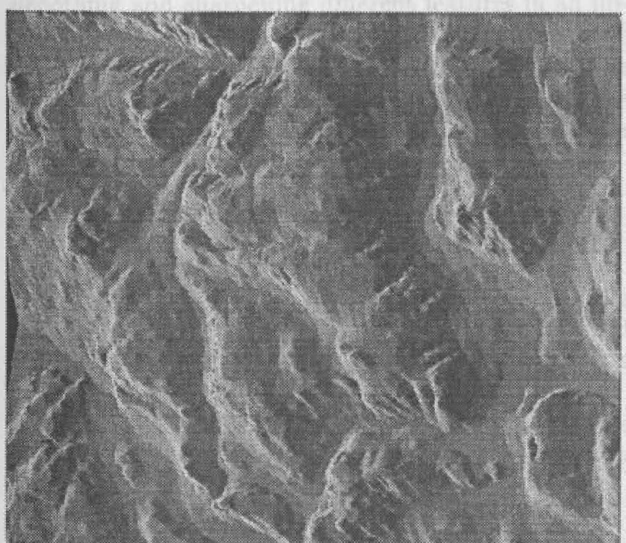


Figure 4: The simulated SAR image generated from the DTED shown in Fig. 3 and the known SAR platform parameters.

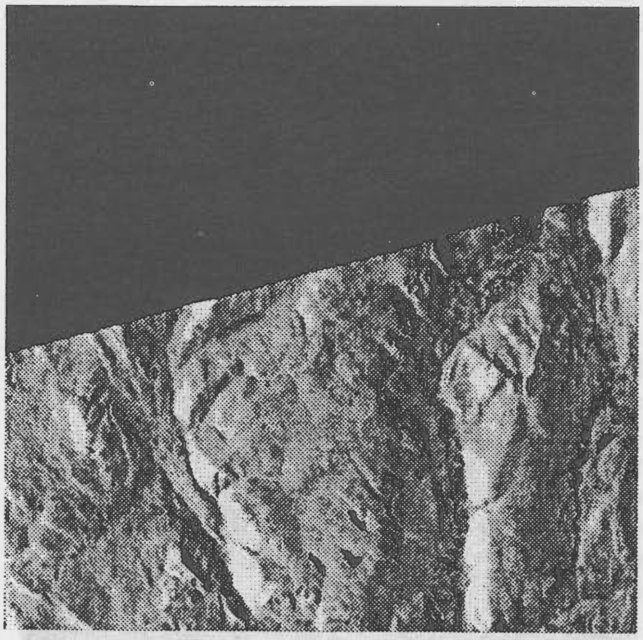


Figure 5: The geocoded image of Fig. 2 to the same lat./long. grid as the DTED image. (©ESA 1991)

Speckle can be modelled as a multiplicative noise process. Both Frost [1] and Lee [5] devised linear adaptive filters which incorporate multiplicative noise statistical properties. Both filters employ a minimum mean square error (MMSE) approach and are computationally efficient. The results of filtering a SAR image (Fig. 6) are shown in Fig. 7 (Frost) and Fig. 8 (Lee). These filters exhibit good speckle reduction with minimal loss of sharpness.

Both the Frost and Lee filters do not assume an explicit model for the underlying (speckle-noiseless) signal and incorporate only its local mean and variance. As well, these MMSE filters assume that the speckle is everywhere fully developed and are optimal only if both the received and underlying signals are gaussian. These shortcomings are addressed using a maximum a posteriori (MAP) filter considered by Kuan et al. [4] and modified by Lopes et al. [6]. The MAP filter implemented in EV-SAR allows modelling of the underlying signal using either a symmetrical Beta or Gamma distribution. It locally determines thresholds, above and below which filtering is not applied, instead retaining the pixel value in question or replacing it by the local mean. The result of applying the MAP filter is shown in Fig. 9. A refinement can also be done by utilizing edge and line ratio detectors to separate texturally different regions in the local application of the MAP filter. Figure 10 shows the results of the MAP-refined filter. The MAP and MAP-refined filters provide excellent results, suitable for subsequent classification of the image contents.

Figure 6: Original SAR image of water (left), ice (centre), and land (right). (SAR image was produced by the CCRS C/X Airborne SAR, courtesy of CCRS, NRCan, Canada.)

Figure 8: Lee filtered SAR image.

Figure 7: Frost filtered SAR image.

Figure 9: MAP filtered SAR image.

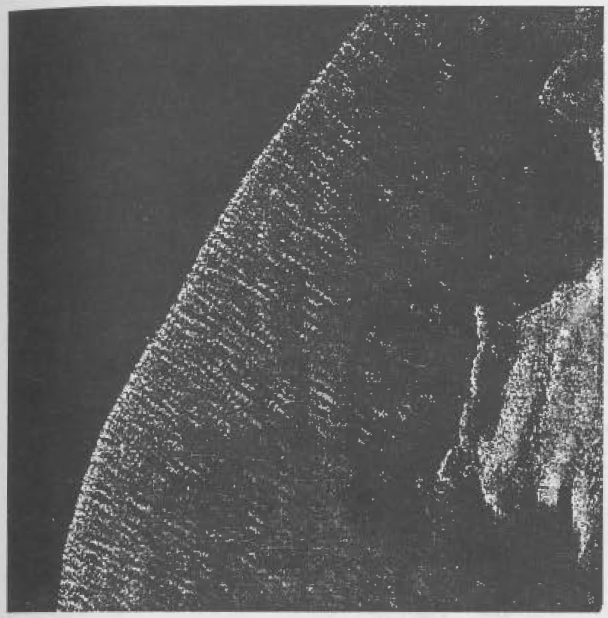


Figure 10: MAP-refined filtered SAR image.

3.2 Image Processing Filters

In addition to the SAR-specific filters in EV-SAR, the standard EarthView commands include more general filtering techniques. Both median filtering and lower-upper-middle (LUM) filtering can be effective in removing speckle and other noise. The LUM filter is a rank ordered filter and is capable of simultaneously smoothing and sharpening different features in an image. As well, EarthView provides spatial and spectral convolution filtering utilizing finite impulse response (FIR) or boxcar techniques and FFTs. These provide the basics for removing or enhancing specific image content.

3.3 Histogram Modification

Image enhancement can be achieved by modifying image data such that it more closely matches a desired distribution. Often it is desirable to stretch a particular region of intensity values to provide more contrast. Also, the expected distribution of an image may be known. EarthView provides the capability to modify image data to more closely match a uniform, cube root, logarithmic, exponential or Rayleigh distribution.

3.4 SAR Image Classification

EarthView provides standard multispectral classification methods including box, minimum-distance and maximum-likelihood. A training function is provided allowing the selection of multiple polygonal regions for multiple class types. Texture based classification can also be performed using EarthView. This classification is based on the texture unit described by He and Wang [3].

4. SAR IMAGE DISPLAY AND MEASUREMENT

SAR data sets are typically very large. One scene of 100km×100km at 25m resolution is 8000×8000 pixels, 16 bits per pixel, for 122 MB of data. The large data sets combined with the high dynamic range of the imagery requires special processing. Since all 8000×8000 pixels cannot be displayed on screen at once, EV-SAR displays images using automatic pixel averaging to low pass filter and resample the image data to conform to the display screen pixel dimensions. Dynamic range of the images are rescaled for the gray scale range of the display. To access the full resolution of the data, interactive image zooming and panning is available. To enhance interpretation of the data values pixel value readoff is possible as is latitude/longitude tagging to find a pixel at a specific geodetic coordinate.

4.1 Complex Image Display

One special type of SAR imagery is complex imagery where each pixel value is a complex number with a magnitude and phase. The phase component contains fine range information. When two SAR images are multiplied pixel by pixel, an interferogram is made containing a fringe pattern which may be phase unwrapped to derive terrain elevation information. It is thus necessary to display the complex image magnitude and phase either separately or combined at once. To optimize the display of complex imagery, EV-SAR maps the complex image magnitude into image brightness and the phase component into a colour wheel of red merging to green, then blue, and finally back to red again completing the circle. Optionally, a phase-only image can be displayed. An example of an interferometric SAR image is given in Figure 11 where the magnitude and phase are shown in two separate images.

4.2 Point Target Response Analysis

Determining the quality of a SAR image is important for correctly interpreting and analyzing the data. This is particularly important when the imagery has been post processed using spacial filters and resampling. The most accepted method for measuring SAR image quality is the analysis of the point target response of a strong point target, such as a calibration reflector, against a radiometrically dark background. Point target measurements in EV-SAR are made by interactively windowing a point target in a displayed image, and then running an analysis on the target. The analysis function interpolates the point target response, and produces a three-dimensional plot of the point target

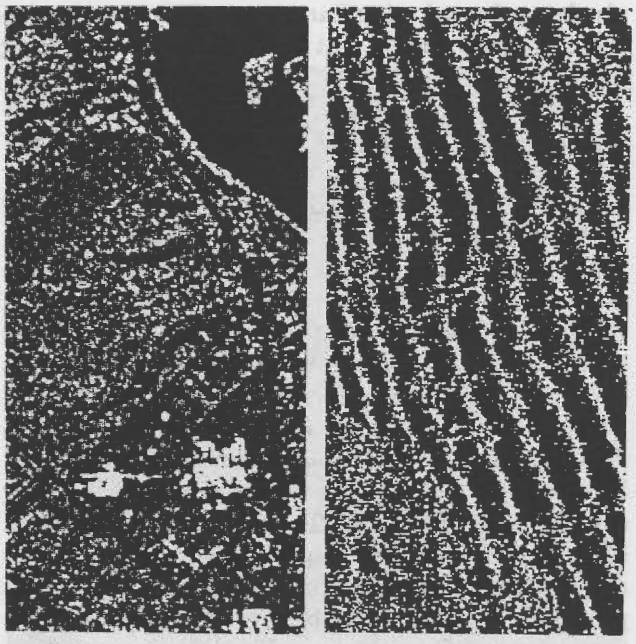


Figure 11: An example of an interferometric SAR image display with the magnitude component (left) and the phase component (right). (SAR image was produced by the CCRS C/X Airborne SAR, courtesy of CCRS, NRCan, Canada.)

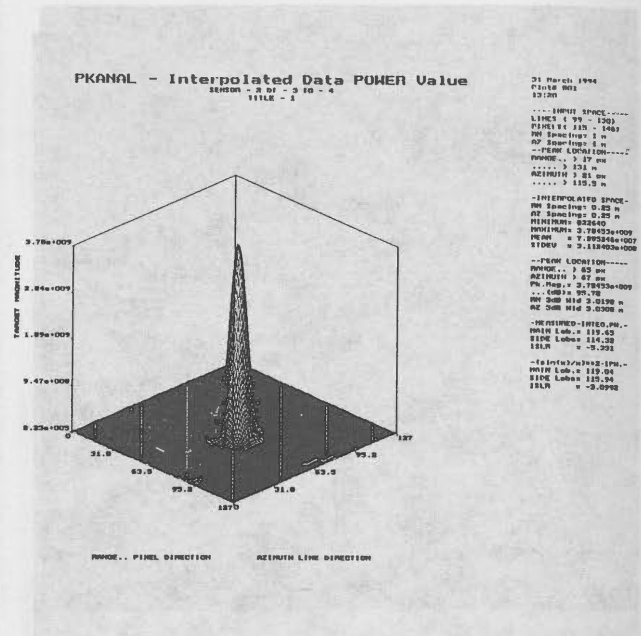


Figure 12: Example output produced by the point target analysis functions.

response as well as a contour plot, a range profile plot, and an azimuth profile plot. These plots are useful for qualitative evaluation of the imagery. In addition, quantitative measurements calculated include the -3dB impulse response (IRW) widths in range and azimuth, integrated sidelobe ratio (ISLR), and peak to sidelobe ratios (PSL). The equivalent number of looks in an image is measured by windowing a uniform area in the image such as a field and comparing the mean and variance statistics for this windowed area. Examples of the point target analysis plots are given in Figure 12.

4.3 Map Projections

Most SAR imagery is produced in the ground range-azimuth coordinates. This coordinate system is convenient for SAR image production at the ground station, but less so for the interpreter. Thus in addition to geocoding, it is possible to reproject an image into a standard map projection such Lambert Conformal Conic (LCC), Universal Transverse Mercator (UTM), or Universal Polar Stereographic (UPS). Once projected, the image may be displayed with full latitude/longitude tagging, pixel value read-off, zoom and pan.

5. POLARIMETRY

Conventional SAR systems transmit and receive electromagnetic waves of a fixed, single polarization. Polarimetric SAR systems transmit two electromagnetic waves of orthogonal polarization (commonly horizontal and vertical) and measure the orthogonal components of each reflected wave. These four measurements can lead to the determination of the scattering matrix for each return cell in a SAR scene. The reflected waves from many surface features have strong dependencies on the polarization of the incident wave. The additional information provided by polarimetric SAR can be used to improve identification and classification of features in SAR imagery.

5.1 Image Synthesis

Through proper calibration, the four polarimetric measurements can be used to derive the scattering matrix for each pixel in a SAR scene. Thus for an antenna transmitting an arbitrarily polarized wave, the polarization of the reflected wave can be determined from the scattering matrix. Finally, given an arbitrarily polarized receive antenna orientation the expected SAR system measurement can be calculated. Performing this calculation for an entire SAR scene results in image synthesis. EV-SAR provides an interactive function for synthesizing images given theoretical SAR systems having arbitrary transmit and receive antenna orientations. Both the receive and transmit orientations can be independently specified and incremented, allowing for an animated sequence of synthesized images.

5.2 Polarization Signatures

A useful tool for analyzing and interpreting polarimet-

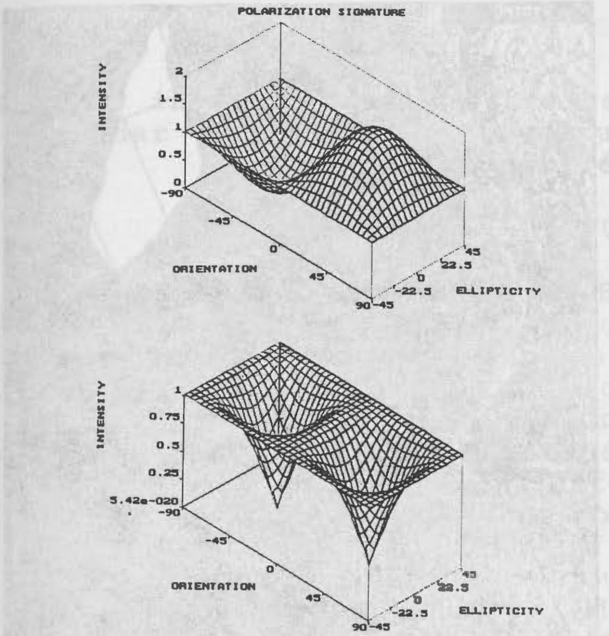


Figure 13: Polarization signatures of an ideal reflector.

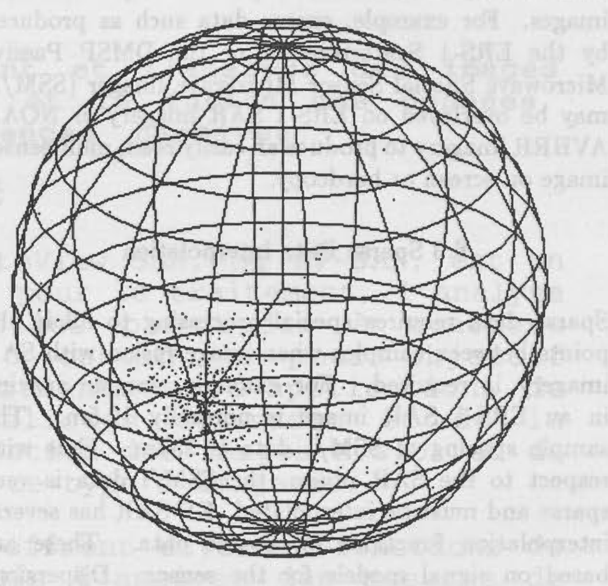


Figure 14: Display of scatterers on the Poincaré sphere.

ric behaviour of specific scatterers is the polarization signature [7]. A polarization signature consists of a three-dimensional plot of backscatter measurement (ultimately the radar cross section) for a particular distributed area as a function of the elipticity and orientation angles of the transmit antenna. The receive antenna can be polarized either the same (copolarized) or orthogonal to (cross polarized) the transmit antenna. Polarization signatures of an ideal reflector as generated by EV-SAR are shown in Figure 13. These signatures are convenient for exhibiting scattering behaviour of differing surface areas.

5.3 Poincaré Sphere

Another useful tool for observing differences in scattering behaviour is the Poincaré sphere. For an arbitrary transmit antenna orientation the polarization of the reflected wave can be characterized by a set of quantities known as Stokes parameters. The Stokes parameters are similar in definition to spherical polar coordinates and therefore lend themselves to a spherical representation. The Stokes parameters of many scatterers can be mapped onto the surface of a sphere, such that a given location corresponds to a particular polarization. EV-SAR provides the ability to interactively map the polarimetric information of chosen scatterers on the surface of the Poincaré sphere (Fig. 14). The transmit antenna polarization can be specified, and the sphere can be rotated to provide a view from any direction.

6. MULTISENSOR DATA FUSION

Since SAR measures the backscattering coefficient at RF frequencies (greater than 1.0 GHz), the imagery is complementary to many other types of remote sensing data. Multisensor data fusion describes the range of techniques that are used to combine imagery from different sensors into one image for interpretation.

6.1 Image Coregistration

For fusion of any two data sets, there must be a mapping between the pixels in one image to those in the other. Image coregistration exploits this mapping to exactly line up the two images so that they overlay perfectly. Two methods of coregistration are possible with EV-SAR. First the images may be coregistered manually using ground control points (Section 2.3). This is very accurate for small scenes where the respective image geometry mapping is approximately linear. Second method is to simply use the latitude/longitude tagging of each image to resample one image (the slave) to be coregistered to the other image (the master). This second method is the only possible solution when the slave image data has a far lower resolution as compared to the master.

6.2 Grid, Vector, and Point Overlay

Multisensor data fusion in EV-SAR allows overlay of various types of non-raster data. Latitude/longitude grids may be overlayed as a geographical reference. Vector data such as coastlines extracted from the Digital Chart of the World may be overlayed for reference and to verify the georeferencing. Other types of vector data such as point data samples of wind velocity vectors, ice motion vectors, or sample data with text annotation may be overlayed on raster images. For example, sparse data such as produced by the ERS-1 Scatterometer or the DMSP Passive Microwave Special Sensor Microwave Imager (SSM/I) may be overlayed on ERS-1 SAR imagery or NOAA AVHRR imagery to produce an easily read, multisensor image on screen or hardcopy.

6.3 Sparse Data Interpolation

Sparse data requires special processing to fill in the points between samples when image fusion with SAR imagery is required. For example, sample spacing in an ERS-1 SAR image is normally 12.5m. The sample spacing of SSM/I data is 25km. Thus with respect to the SAR image, the SSM/I data is very sparse and must be interpolated. EV-SAR has several interpolation functions for sparse data. These are based on signal models for the sensor. Dispersion-based techniques use a resolution model for the sensor to interpolate the data points. Interpolation techniques use a spatial model of the data to interpolate points on a surface.

Sparse data interpolation is done in three steps. First, a blank image (the slave image) is filled with zero pixel values and is coregistered to the SAR (master) image. Second, the sparse data points read from a tabular or vector file are placed in the blank image using the lat./long. georeferencing to map the data samples into pixel coordinates. Finally, the third step is the dispersion or interpolation to fill in all the pixels in the slave image.

6.4 RGB and IHS Image Fusion

Image-Image fusion may done in EV-SAR using standard Reg-Green-Blue (RGB) and Intensity-Hue-Saturation (IHS) multichannel image combination techniques. One application of IHS image fusion is the combination of ice concentration data (derived from the DMSP SSM/I passive microwave data) with a SAR image showing in detail the ice flows[8]. A total of three categories of ice concentrations (thin ice, first year ice, and old ice) are coregistered to the ERS-1 SAR image, sparse data interpolated, and converted to IHS. The hue and saturation channels then contain the ice concentration information. The intensity channel is replaced by the ERS-1 SAR image, and the IHS data set is then converted back to RGB for display. The result is an ERS-1 SAR image coloured to represent green for thin ice, blue for first year ice, and red for old ice. The more saturated the colour, the greater the ice concentration. Sample point data and coast lines may be overlayed completing the data product. An example of an ERS-1 SAR image with sample data overlay and latitude/longitude grid is given in Figure 15.

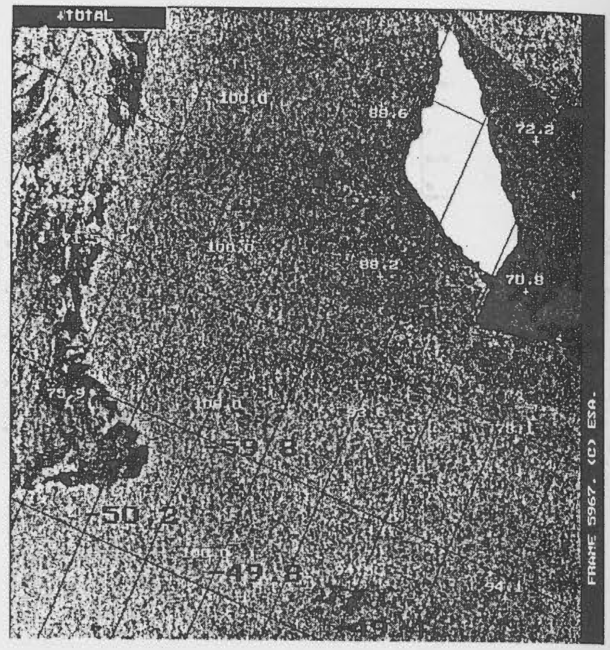


Figure 15: Example ERS-1 SAR image with sample data overlay (Copyright ESA 1992).

7. CONCLUSIONS

Users have found the EV-SAR application package to be a powerful low-cost solution for working with SAR imagery. With the recent availability of the new satellite SAR imagery (ERS-1 1991, JERS-1 1992, RADARSAT 1995), interest in SAR will increase. This package acts as a tutorial in SAR for users who are new to SAR remote sensing. It serves as a tool to aid the expert user in SAR image measurement and interpretation.

ACKNOWLEDGEMENTS
The authors acknowledge the support of the Canada Centre for Remote Sensing, National Resources Canada, and the Ice Services Branch, Atmospheric Environment Service, Environment Canada.

REFERENCES

[1] Frost, V.S, et al., 1982. A model for radar images and its application to adaptive digital filtering of multiplicative noise, IEEE Tr. on PAMI, Vol. 4, No. 2, pp. 157-166, March 1982.

[2] Guindon, B., M. Adair, 1992. Analytic formulation for spaceborne SAR image geocoding and "value-added" product generation using digital elevation data, Canadian Journal of Remote Sensing, 18(1):2–12.

[3] He, D.C., Wang, L., 1989. Texture unit, texture spectrum and texture analysis, IGARRS'89, Vol. 5, pp. 2769-2772.

[4] Kuan, D.T., et al., 1987. Adaptive restoration of images with speckle, IEEE ASSP, Vol. 35, No. 3, pp. 373-383, March 1987.

[5] Lee, Jong-Sen, 1981. Speckle analysis and smoothing of synthetic aperture radar images, Computer Graphics and Image Processing, Vol. 17, pp. 24-32, 1981.

[6] Lopes, A., et al., 1990. Maximum a posteriori speckle filtering and first order texture models in SAR images, Proceedings of IGARRS'90, Vol. III, pp. 2409-2412, Washington D.C., May 1990.

[7] Zebker, H.A., Van Zyl, J.J., 1991. Imaging radar polarimetry: a review, Proceedings of the IEEE, Vol. 79, No. 11, November 1991.

[8] Ramseier, Reneé O., Andrew Emmons, Bernard Armour, Caren Garrity, 1993. Fusion of ERS-1 SAR and SSM/I Ice Data. Proceedings of the 2nd ERS-1 Symposium, Hamberg Germany, October 11–14 1993.

Système intégré pour le traitement et l'analyse des images de radar à synthèse d'ouverture et la fusion des données radar et hyperfréquences passives

RÉSUMÉ

Le progiciel d'application EarthView SAR, ou *EV-SAR*, est un ensemble intégré de logiciels conçus pour le traitement, l'analyse et la fusion de données recueillies par radar à synthèse d'ouverture (R.S.O.). Ce projet continu a pour objectif l'élaboration d'un ensemble d'instruments et d'applications de pointe visant à aider les spécialistes du radar et de la télédétection à extraire de l'information des images R.S.O. et de créer de nouvelles applications pour ce type d'imagerie.

Les logiciels créés à ce jour offrent diverses fonctions de traitement des images R.S.O., depuis l'analyse des mesures de qualité de l'image jusqu'à la fusion d'images R.S.O. avec d'autres ensembles de données. Ces fonctions sont les suivantes :

1. correction de diagramme de rayonnement et conversion distance oblique-distance au sol;
2. simulation R.S.O. et géocodage;
3. filtrage d'images R.S.O.;
4. techniques d'affichage d'images complexes;
5. techniques d'analyse polarimétrique;
6. analyse de réponse de cible ponctuelle;
7. R.S.O./DMSP-SSMI/diffusiomètre ERS-1/fusion des données des calques vectoriels.

Pour effectuer des corrections radiométriques, on détermine le diagramme d'antenne approximatif et on élimine l'effet des gains d'antenne non uniformes sur les images R.S.O. Les distorsions des images spatiales attribuables à la géométrie de l'imagerie R.S.O. et aux altitudes de terrain sont corrigées au moyen de fonctions permettant d'effectuer la conversion distance oblique-distance au sol des images et le géocodage des images de terrains accidentés. Les filtres utilisés avec les images R.S.O. sont conçus spécialement pour corriger le bruit multiplicatif des données R.S.O. On a élaboré des techniques d'affichage d'images complexes permettant d'indiquer les phases par des couleurs. Plusieurs fonctions sont utilisées pour l'analyse de l'imagerie R.S.O. polarimétrique, notamment la synthèse polarimétrique et l'affichage de données sur une sphère de Poincaré. Enfin, on a élaboré un ensemble de fonctions de fusion de données permettant d'intégrer des ensembles de données de diffusiomètre radar et hyperfréquences passives avec des images R.S.O. à l'aide des techniques matricielles et vectorielles.

L'élaboration d'applications pour l'imagerie R.S.O. et l'exploitation de cette technique représentent un défi perpétuel. Avec l'arrivée du RADARSAT, les participants à ce projet fournissent un effort soutenu afin d'élaborer des techniques d'interprétation des images R.S.O. utiles et innovatrices.

PROSPECT OF HIGH RESOLUTION COLOUR IMAGERY IN NEW BRUNSWICK

RÉJEAN H. CASTONGUAY

New Brunswick Geographic Information Corporation
985 College Hill Rd, Fredericton, N.B., E3B 5H1, Canada

RONALD ROBICHAUD

GEOMACADIE Services Ltd.
172 Saint-Patrick Street, Bathurst, N.B., E2A 1C7, Canada

JEAN-PIERRE ANGERS

Université de Moncton
Centre universitaire Saint-Louis-Maillet
165 Hébert Boulevard, Edmundston, N.B., E3V 2S8, Canada

KEY WORDS: Soft Orthophoto, CARIS, GIS, DEM, Digital Orthophoto

ABSTRACT

In 1970, New Brunswick began production of a series of orthophoto graphic maps at a scale of 1:10 000. Those series were completed in 1983. Orthophoto maps, produced in three series (planimetric, topographic and cadastral) were very popular then and remain so, even though the maps are becoming somewhat outdated.

In 1983, New Brunswick undertook the creation of a digital topographic data base to 1:10 000 standards. This data base, including a digital elevation model (DEM), will be completed in 1994.

Orthoimages continue to be an attractive solution in the management of New Brunswick's land mass. There is a new enthusiasm for the soft orthoimage, and a number of projects are under way involving collaboration between the universities, the geomatics industry and the New Brunswick Geographic Information Corporation. This paper looks at one of the projects recently initiated. Moreover, there is a possibility that the provincial data base revision will be done on the basis of orthoimages. Soft orthoimages could also complement the digital topographic data base in the year 2000.

RÉSUMÉ

En 1970, le Nouveau-Brunswick se lançait dans la création d'une série de cartes orthophotographiques au 1:10 000. Cette série fut completée en 1983. L'orthophoto, produite en trois séries, c'est-à-dire sous forme planimétrique, topographique et cadastrale fut et demeure encore très populaire même si les cartes commencent à être un peu dépassées.

En 1983, le Nouveau-Brunswick entreprenait la création d'une base de données topographiques numériques selon les normes du 1:10 000. Cette base de données, laquelle inclut un modèle numérique de terrain (MNT), sera achevée en 1994.

Les orthoimages demeurent une solution attrayante dans la gestion du territoire au Nouveau-Brunwick. L'orthoimage électronique crée un nouvel engouement et plusieurs projets de collaboration entre les universités, l'industrie de la géomatique et la Corporation d'information géographique du Nouveau-Brunswick sont en cours. Cet exposé décrit un des projets récemment entrepris. Notons que la révision de la base de données provinciale pourrait possiblement être effectuée à partir d'orthoimages. Ces dernières pourraient aussi être un complément à la base de données topographiques numériques.

1. INTRODUCTION

Since the 1950's, the urban and resource territories of New Brunswick have been mapped and remapped, and for each series of maps, new technologies were used. Today, urban maps are mostly hard copy and they are aging rapidly. As well we are approaching the completion of the third resource series. We are asking questions about what our infostructure needs and about the economical way of meeting these needs. On the one hand, we have invested over $10 million in a soft 1:10 000 infostructure we would like to preserve; and on the other, we are facing the demand from the new "generation": the video generation and the image management generation. The soft colour orthoimagery (SCO) represents perhaps an economically justifiable product because it can be used to protect the last decade's investment, and in fact, provide management tools for the future.

This article describes one of three (3) New Brunswick projects that evaluates the soft colour orthoimagery as a product for providing urban and resource infostructure and revising New Brunswick's digital topographic database at 1:10 000 and therefore offering us a glimpse into the future.

2. MAPPING HISTORY

In New Brunswick the need for an infostructure for the collection of geographical information was identified early in the establishment of the colony. However, it is under Thomas Baillie (1824-1851) the Surveyor-General of New Brunswick, that it was most stressed. This is illustrated when in 1838, the lieutenant governor, John Harvey, requested Baillie to show maps of the province to Charles Butler, a member of Lord Durham's commission charged with increasing the general prosperity and wealth of the colonies. Baillies reply was

> "...that unless these Surveys are scientifically and carefully connected on the ground and complied on

paper it will be impossible for me to lay before Mr. Butler such plans of the Province as will exhibit the vacant crown lands with any degree of accuracy and this deficiency of information will be the more apparent as Upper Canada has been carefully surveyed and very excellent compilations made therefrom".

Ten years later, in 1848 little progress had been made and Baillie was quite desperate about the system. In a letter to Privy Council he had this to say

"...the present system, if system it can be called, of surveying isolated lots of land by separate and often ill connected surveys, has presented to my mind only a complete mass of confusion, but having repeatedly represented the evil effects which must at some future period inevitably ensure, I felt that I had done my duty".

2.1 The "20 Chain" Map Series

It was nearly one hundred years later (1944) in the Brief on "Forestry and Post-war Reconstitution in New Brunswick" (Burchill 1944) that the need for a province-wide map series was recommended and acted upon. It had this to say:

"WHAT MUST BE DONE

To take advantage of the "opportunity" and remedy the faults of the "present situation," action is necessary by both government and industry. The plan of action which we recommend includes certain projects which should be carried out as soon as men and equipment are available.

Among these are:

1. Aerial photographs and maps of the whole province..."

Soon after the report was published, the project to map the whole territory was begun, and 10 years later, for the first time in history, the province had a series of maps at the scale of 1 inch = 20 chains.

2.2 The Orthophoto Series

Toward the end of the 1960's, the series of post-war maps could no longer meet the New Brunswick needs. It had to be revised. Even though the best technology of the day had been used, its accuracy no longer corresponded with current needs. This was partly because of the imprecise control survey network and the rather unsophisticated methods of mechanical adjustment and compilation of topographic features.

From 1960 to 1963, the primary control survey network was densified (Hamilton 1974), and from 1969 to 1972, a secondary control survey network was established. Aerotriangulation, adjustment, and photogrammetric methods had developed considerably. A decision was made, therefore, to produce a new base map series instead of revising the series of post-war maps.

After reflecting on the matter for several years, in 1971, the decision was made to go ahead with an orthophotographic base map at the scale of 1:10 000. Three versions were produced:

1) the planimetric maps, to which were added certain themes such as hydrography through the use of photogrammetry, symbols representing monuments, etc.;
2) the hypsographic maps with 5 metre contour lines; and
3) the property maps.

It should not be surprising that the orthophoto was chosen as the basic map. One need simply to remember that, toward the end of the 1960's, Canadian social democracy had reached its peak. It was probably the first time in Canada that insistence was so firm on the importance of acquiring a tool, being the orthophoto, that the public could understand, a tool that would enable us to communicate with it, as well as encourage the public to participate in and contribute to the decision-making process. It was the backdrop against which the link between the public and spatial information was established. Implementation of the orthophoto meant more than recognizing the public's right to spatial information; it meant giving it the tools of access to information, offering it directly products that everyone could comprehend. The orthophoto's success was extraordinary (Castonguay 1984).

Although the public was enchanted by the orthophoto, in the 1980's in most cases Canadian administrators of geomatics programs shrank back from the social viewpoint and encouraged a return to standard map making.

During the 1970's and early 1980's, the digital representation became more and more popular. Digital topographical conventional maps were sold with the promise that revision would be economical and that it would help us to handle all kinds of spatial questions. The orthophoto, even though digitally produced, did not lend itself well to revision. The decision was therefore to produce a new base map series.

2.3 The Digital Topographic Base Series

In 1983, New Brunswick embarked on a program to create a digital topographic base according to a standard equivalent to the production of maps at the scale of 1:10 000. The data base is to be completed in 1994. But now, in some regions (Figure 2-1), this data base is already 10 years old. Unless the data is updated, the investment of over $10 million will lose more and more of its value.

BLOCKS OF AERIAL PHOTOGRAPHY

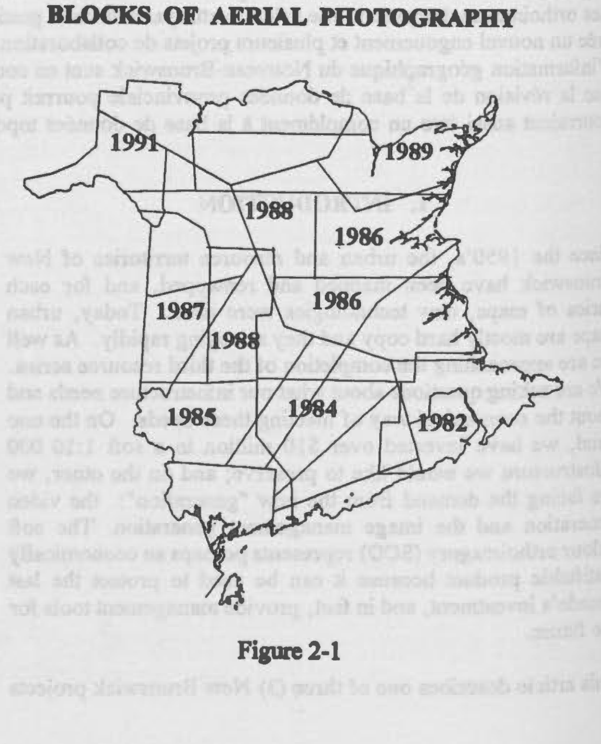


Figure 2-1

As indicated, each time the basic map was revised, the province began from scratch with new aerial photography, aerotriangulation, etc. Consequently, little experience in updating was acquired over the years.

3. SPATIAL INFOSTRUCTURE NEEDS AND CONSTRAINTS

The government of New Brunswick recognized the importance of information about land, water and related resources to manage these valuable resources. In order to establish a coherence to the government's efforts as they relate to land information management, in 1989 a Land Information Policy was established.

This policy ensures that the collective efforts best support the provinces' need for geographically related information to set policies and regulations for land and water use, and for the resource management. The Policy establishes the Geographic Information Corporations digital 1:10 000 topographic data base as the official base on which all other provincial spatially referenced data will be founded. This collective data base will have to be updated if it is to maintain its usefulness. How will it be revised? There are many options. However, funds for maintaining collective projects are shrinking. In 1992, the Geographic Information Corporation adopted a Pricing Policy for its products and services. In the category of existing basic products and services that involve only maintenance, updating, and marketing expenses, the pricing method is based on the recovery of additional costs. In other words, costs for revision must not exceed revenues from the digital topographic data base. We are therefore obliged to find more cost effective methods for revising the infostructure.

The cooperative model (Hamilton 1976) has always been appealing. Given that there are only a few province-wide GIS users/producers, it will take some time yet for this model to reach full maturity. Furthermore, as demonstrated in Figure 2-1, the digital data of the digital topographic data base is already over 10 years old in some areas. Some catch-up techniques will perhaps have to be devised to update the digital topographic data base. In that light a project was undertaken to assess the soft colour orthophoto (SCO) product for updating the digital topographic data base.

4. CONCEPT

The production of most soft orthoimagery and the subsequent revision of digital topographic data is essentially the same as the traditional mapping approach. In summary it consists of the establishment of geodetic control, aerial photography, photo control, aerotriangulation, block adjustment, collection of digital terrain data, orthophoto production and interactive revision. This approach does not lend itself to significant savings. However, the concept described below bypasses many traditional phases.

4.1 Geodetic Control

In New Brunswick over 15 000 geodetic control points have been established between the late 60's and the mid 70's. Consequently, there is no need for additional control.

4.2 Aerial Photographs

In 1992, the Department of Natural Resources and Energy undertook a study to decide on the methodology to be used for updating the forest inventory. In the first digital forest inventory, deciduous species were not included. Because of the resurgence of interest in these species, there is a need for them to be inventoried in the next forest inventory. In order to study the new method of inventorying, a sector of the province that contained all arboricultural species was photographed. The pilot project's aerial photographs included colour photography at the scale of 1:30 000.

The results of the study were conclusive and the whole province will be photographed with colour photography at the scale of 1:25 000.

This aerial photography for the project will be taken in the fall, before leaves fell. If the infostructure is to be revised from aerial photography the ideal time for taking photographs would be before the leaves appear or after they have fallen. However, because colour photography will be available for the whole province, the use of this photography, if satisfactory, would entail no additional aerial photography cost.

4.3 Photo Control, Aerotriangulation and Adjustment

The purpose of photo control, aerotriangulation and adjustment is to establish control to carry out the relative and absolute orientations of each photographic model. As discussed earlier, the digital topographic data base will be completed in 1994. The question raised was as follows: Can the digital topographic data be used as control to generate the soft colour orthophoto? If this was successful, costs related to photo control, aerotriangulation and adjustment would be eliminated.

4.4 Digital Elevation Model (DEM)

During the creation of the digital topographic database, spot elevations were collected in an irregular grid, 1 mm to 3 mm at photo scale, (1:35 000 photo) depending on the roughness of the terrain. Block adjustments were carried out as per the 1:10 000 map accuracy specifications and the spot heights were read to a 0.1 m precision.

Usually the creation of the DEM is an expensive component in the orthophoto process. In New Brunswick the DEM collection is near completion. The question now is: Will the investment in DEM collection over the last ten years allow the province to update the digital topographic base at a reduced cost?

5. EVALUATION OF THE CONCEPT

5.1 Partnership

The contract to evaluate the concept was awarded to Universal Systems Limited (USL) of Fredericton. USL worked in close collaboration with Northeast Exploration Services Ltd. and Geomacadie Services Ltd. Northeast Exploration was given the orthophoto production responsibilities and Geomacadie the map revision responsibility.

5.2 Selection of Imagery and its Quality

Aerial photography flown to produce vector maps can vary in quality, within the range of specifications, without having a great impact on the data. However, aerial photography for soft orthophoto is much less forgiving. The quality of the aerial photography used in the project was not the best to produce an orthophoto. There was a lot of haze and the shadows were at their maximum. The shadows are large because the photograph was taken a few minutes from the limit of the minimum solar angle allowable. It is good, however, to be confronted with the worst conditions possible because this helps us to appreciate all the problems associated with updating topographic data using the SCO.

5.3 Software Selection

In 1983, New Brunswick chose the CARIS software to produce its digital topographic infostructure. Since then, CARIS software developments (Derenyi 1991)(Derenyi and Teng 1992) added the soft orthophoto functionality. Subsequently Universal Systems Ltd. incorporated and enhanced the software development as a standard package in the CARIS products (Masry, Mayer and Clarke 1994). Although the first objective of the project was to evaluate the soft orthophoto technology as a tool for digital data revision, it was also an opportunity to gain some experience with the orthorectification module of the CARIS software.

5.4 Hardware Selection

In a large project, one may select a powerful computer system to optimize the human and hardware resources. In an evaluation project, the hardware resources are not critical. Both Northeast Exploration Services and Geomacadie Services used a similar SUN SPARC Station IPX, equipped with a 2.6 GB hard disk, a 2.3 GB tape drive, a 16 MB RAM, and a 16-inch standard colour monitor.

EXPERIMENTAL SITE

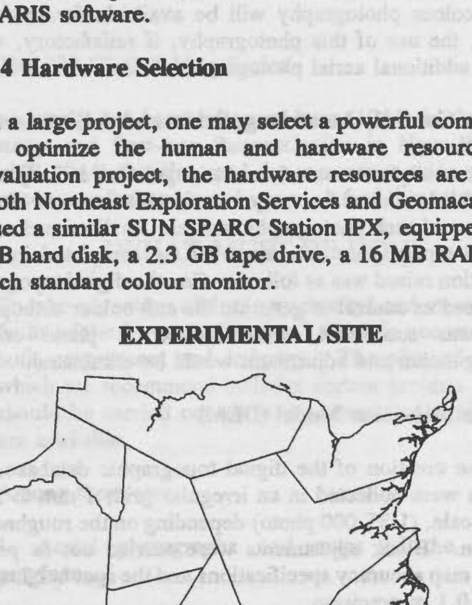


Figure 5-1

5.5 Spatial Extent

As described earlier, the philosophy is to use existing photography to minimize the revision cost of the digital topographic base. The existing photography was flown at the experimental forest site. The Department of Natural Resources and Energy's experimental forest site is located in Central New Brunswick (see Figure 5-1). The extent of the site is equivalent to a map at a scale of 1:50 000. The New Brunswick topographic system is a breakdown of the 1:50 000 NTS. The 1:50 000 encompasses twenty-five maps at a scale of 1:10 000. Therefore, the aerial photography covered twenty-five 1:10 000 maps. The file for the Village of Stanley was the one chosen because the area provides the most changes in topographic feature and height variations (160 metres), a very good area for checking the orthophotographic and mosaicking processes. The aerial photography lines fall near the northern and southern edges of the map (see Figure 5-2), which enables us to check the mosaicking in the direction of the flight lines as well as between them. Also, the lateral mosaicking gives us a clearer idea of the software's efficiency in handling radiometric corrections between the flight lines.

FLIGHT LINES

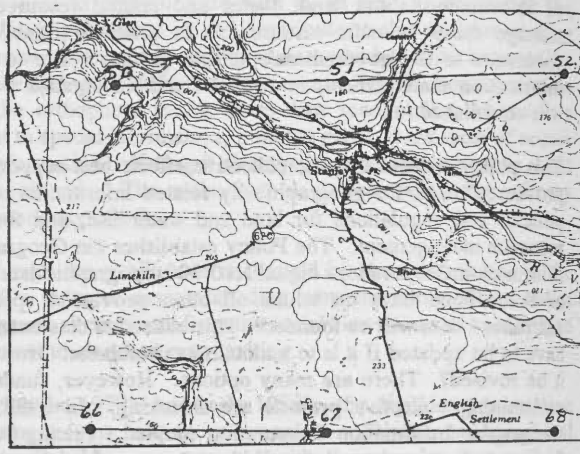


Figure 5-2

5.6 Phases

5.6.1 Input

Three intrinsic data bases (Figure 5-3) are essential for execution of the project: digital aerial photography (DAP), digital topographic data (DTD) and digital elevation models (DEM).

A large number of scanners in the industry are capable of scanning aerial photography. The cost of such scanners varies considerably. It is very likely that the geometric and optical qualities of these scanners are generally in proportion to their cost. Quite possibly, there are very good scanners at a reasonable price, and it is also quite possible to perform geometric calibrations on a scanner. In a production environment, the choice of an appropriate scanner is essential. In our situation, the project's variables had to be minimized. For this reason, it was decided to do the scanning on a precision scanner (Zeiss PhotoScan PS 1).

The film (negative) was sent to the company responsible for the scanning (Triathalon mapping Services) so as to be sure that the diapositives could be printed at the optimal density for scanning, that they would be free of ink or grease spots, fingerprints, scratches, and that the emulsion would be on the proper side of the photographic medium.

The scanning was done at a resolution of 15 μm, which corresponds to a resolution of 0.45 metre on the ground.

The use of 24 bits in the scanning made it possible to obtain 16.7 million colours. The data was provided on an Exabyte-type magnetic medium in the Tag Image File Format (TIFF).

The first reason for these specifications was to capture all the details contained in the aerial photograph digitally. The second reason was to provide scanned data appropriate for two projects. The first project is the one discussed in this paper and the second one is to explore the potential application of the high resolution soft colour orthophoto for the revision of the forest inventory.

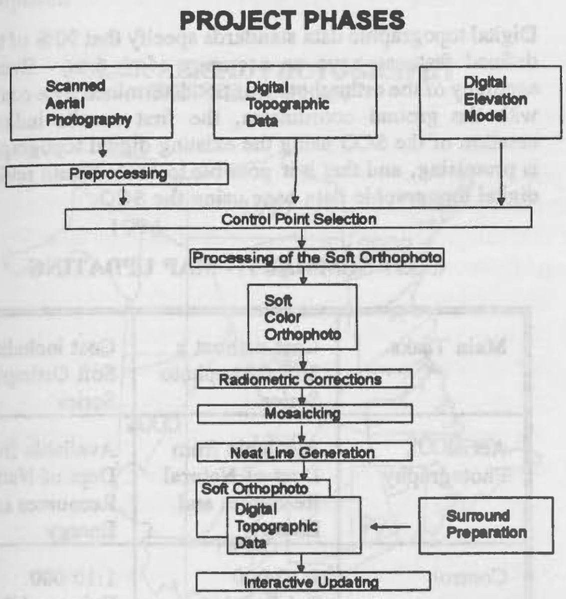


Figure 5-3

The Geographic Information Corporation provided the contractors with the digital topographic data and the digital elevation models. To ensure proper rectification control, nine DTD and DEM files, a 3x3 file matrix, were provided to the contractors in the CARIS NTX format.

5.6.2 Preprocessing

The files produced by scanning at a resolution of 15 μm were approximately 700 MB. The SUN SPARC station IPX assigned to the project cannot use the CARIS software in its 24-bit version. The scanned data had to be preprocessed in order to convert the data from 24 to 8 bits. The preprocessing reduced the size of each photo's files from 700 MB (theoretical size) to about 235 MB (actual size).

The scanned data received were in accordance with the Red, Green and Blue (RGB) model, which makes it possible to define over 16.7 million colours. However, the data was transferred from the RGB model to the Intensity, Hue and Saturation (IHS) or pseudo-colour model. The change in model imposed severe restrictions on colour manipulation.

5.6.3 Selection of Control Points

The first step consisted of displaying the digital topographic data and the uncorrected scanned photo, using the CARIS's editor module (CARED).

The second step was to digitize the coordinates of four fiducials marks (corner or middle) using the zoom function.

The last step involved identifying common points on the scanned photo and in the digital topographic data base by displaying the two data sets in adjacent windows. Five common control points (one near each corner and one near the centre of the photo) are sufficient for the orthophoto transformation. Choosing ten, fifteen, or twenty control points does not improve the results appreciably.

5.6.4 Parameters of the Photogrammetric Camera

Certain data, such as the focal length and the principal point (centre point) of the photogrammetric camera, are also necessary. They were recorded in a file used by the software to resolve the orthophoto transformation.

5.6.5 Production of the Soft Orthophoto

The transformation module requires several data entries and lends itself well to batch processing. The module asks for the names, list of files needed, and the names of the files to be created.

Input Files:

- Name of scanned photo
- Name of DEM data file
- Name of control points file (Recorded from the topographic data base)
- Name of control points file (Recorded from scanned aerial photography)
- Name of fiducial marks file
- Name of photogrammetric camera parameters file

Output File:

- Name of orthophoto data file

Other parameters required:

- Resolution of orthophoto data
- Resolution of output files must be specified. In order to limit the size of the output files a 1.0 m resolution was chosen
- Type of transformation (e.g. orthophoto)
- Type of pixel selection (e.g., nearest neighbour)
- Sector of the photo to be transformed; that is a selected portion or the whole photo

5.6.6 Radiometric Corrections

The purpose of radiometric corrections is to make the appearance of the image in one photo uniform with that of another in the direction of the flight line and between the flight lines. Reduction of the data from 24 bits to 8 bits and the change from the RGB model to the IHS model imposed significant restrictions on the radiometric corrections during the project. The only corrections possible were choosing a palette of colours for one photo and adjusting the other photos to this palette.

5.6.7 Mosaicking

Mosaicking consisted of assembling overlapping photos utilizing the various methods available in the software. The method chosen was to digitize a line on the monitor, digitizing on it as many points as are needed to reduce evidence of the seam as much as possible. As in conventional mosaics, the choice of a line along an uninterrupted line feature, such as a road, a river, or a power line, produced better results. It is possible to create a transition buffer zone along the joint, but this option does not give satisfactory results when the pseudo-colour model is used.

5.6.8 Neat Line Generation

The mosaicking resulted in a file combining six soft orthophotos, the limits of which went beyond the limits of the files of the digital topographic data on the scale of 1:10 000. By specifying the coordinates of the map corners, a new SCO file was delimited, the limit of which correspond to those of the digital topographic data file.

5.6.9 Interactive Update

The first step was to overlay the soft orthophoto and the topographic data. Because the SCO (1992) is more recent than the topographic data (1988), the planimetric changes were identified by doing a visual scan of the combined images. The zoom function was used as needed. To avoid missing a portion of the map in the revision process a grid of 1 km^2 was used as a template for visually scanning with automatically controlled movement to the right, the left, up and down made systematic revision possible.

The erased data was stored as a historical theme, and the new data was assigned an identification code. It is therefore feasible to display both the old data file and the new data file on the monitor.

An identification code was used for the updates. Since CARIS offers a user number code of twelve (12) characters, the new data has the following code:

GSL C OR 92 09 MD

Cartographer's initials (Marcel Doiron)

Date of photography

Revision Method (Orthophoto)

Digitization Method

Company (Geomacadie Services Limited)

5.6.10 Final Presentation

In order to present an orthophoto prototype that meets the provincial standards of map-type presentations, a surround was created and combined with the SCO and the topographic data. For this, the tool box functions of the map editing module (CARED) were used. In addition to the information provided in the conventional surround, the names of topographic features were accompanied by examples in the form of mini-images cut from the soft orthophoto.

6. RESULTS

The preliminary cost estimate for each stage of the process of updating the digital topographic data is shown in Table 6-1. It shows the cost to update the digital topographic files with and without a soft colour orthophoto series.

To check the accuracy of the soft colour orthophoto in comparison with the digital topographic data, sample features were chosen in all parts of the map. The coordinates were measured in the soft colour orthoimage and the coordinates of the same point measured in the digital topographic data base. The results reflect a mean square error of ± 3.6 m. We should point out that it was only possible to identify well-defined points in the village of Stanley. Outside the village (80% of the points) highway, railway intersections, etc. were chosen.

Digital topographic data standards specify that 90% of the well-defined features have an accuracy of ± 4 m. Even if the accuracy of the orthophoto was not determined by a comparison with the ground coordinates, the first results indicate that creation of the SCO using the existing digital topographic data is promising, and that it is possible to contemplate revising the digital topographic data base using the SCO.

COST SUMMARY - MAP UPDATING

Main Tasks	Cost without a Soft Orthophoto Series	Cost including a Soft Orthophoto Series
Aerial Photography	Available from Dept of Natural Resources and Energy	Available from Dept of Natural Resources and Energy
Control	1:10 000 Enhanced Digital Database	1:10 000 Enhanced Digital Database
Scanning	$ 450	$ 450
Data Input and Rectification	$ 400	$ 400
Mosaicking	-	$ 250
Updating	$ 300	$ 300
Surround	Not Required	Not Required
Quality Control	$ 80	$ 150
Field Check	Extra	Extra
Digital File Management and Archieving	$ 80	$ 150
Project Management	$ 90	$ 150
TOTAL	$ 1 400	$ 1 850

Table 6-1

7. FUTURE

The main objective of the project was to evaluate the soft orthophoto technology as a product to update the digital topographic data base. A better understanding of the process and an approximate revision costs were gained.

In New Brunswick the next logical step toward soft orthophoto and digital topographic data revision is the production of a 1:50 000 block (25 files). This will allow a refinement of the process and the revision cost.

The New Brunswick Department of Natural Resources and Energy has embarked on a ten year colour aerial photography project at the scale of 1:25 000. (Figure 7-1). By the end of 1994, the whole southern part of the province will be available. This will give the province a wide range of sites to refine the cost of the digital topographic data revisions using soft orthophoto.

COLOR AERIAL PHOTOGRAPHY
TEN YEAR PLAN

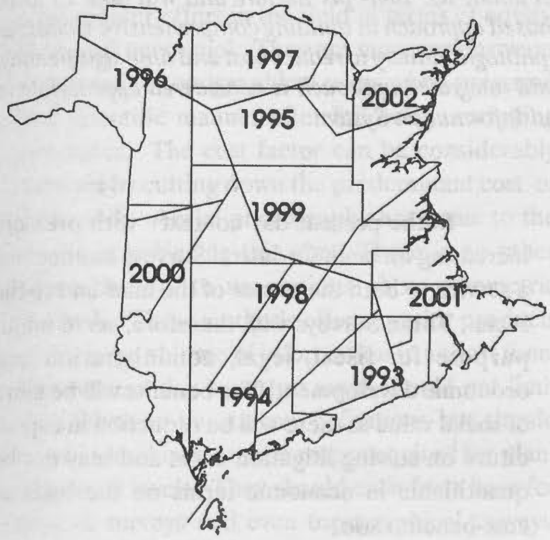


Figure 7-1

Potential cost optimization could be achieved in the following areas:

- optimization of the flight line
- changing the GPS position of the photogrammetric camera from ± 5 m (1993 photography contract) to ± 20 cm
- Scanning costs for use of the Zeiss PhotoScan PS 1 belonging to the Triathalon Mapping Services company in British Columbia were $190 (1993) per colour photo (24 bits) at a resolution of 15 μm. Larger contracts might reduce the cost.
- Production of the soft orthophoto required three (3) hours of processing (each photo) on a SUN SPARC station IPX equipped with 16 MB of central memory. The processing time is reduced to 20 minutes on a HP 9715/50 workstation. Processing time could be reduced to 10 minutes in a production environment.

8. CONCLUSION

The project to evaluate the soft colour orthophoto for the digital topographic data revision has given positive results.

We anticipate the demand for a soft colour orthophoto base will grow in the years to come as an ideal backdrop for parcel mapping, and a host of other applications. The soft colour orthophoto base can become a profitable product on the Electronic Information Highway.

References

Burchill, G.P., and W.E. Golding, 1944. Forestry and Post-War Reconstruction in New Brunswick. New Brunswick Committee on Reconstruction, 35 p.

Castonguay, R.H., 1984. La cartographie en l'an 2000 [Mapping in the Twenty-First Century]. *The Canadian Surveyor*, Vol. 34, No. 3.

Derenyi, E.E., 1991. Design and Development of a Heterogeneous GIS. *CISM Journal*, Vol. 45, No. 4.

Derenyi, E.E. and C.H. Teng, 1992. Digital Map Revision in a Hybrid Geographic Information System, *International Archives of Photogrammetry*, Vol. 29, Part B4, p 533-536, August 2-14, International Society of Photogrammetry and Remote Sensing.

Hamilton, A.C., 1974. A Critical Review of Alternatives with Respect to the Geodetic System for the Maritime Provinces, UNB Department of Geodesy and Geomatics Engineering.

Hamilton, A.C., 1976. The Principal Concepts for a Long-Term Mapping Program in the Maritime Provinces, UNB Department of Geodesy and Geomatics Engineering.

Masry, S.E., L. Mayer and J. Hughes Clarke, 1994. Integrating Geographic Information and Image Analysis Functionality into a Single Spatial Analysis System, *Schriftenreihe des Studiengangs Vermessurgswesen*, Universität der Bundeswehr München, Munich, Germany

New Brunswick Information Geographic Corporation, 1992. Pricing Policy, 4 p.

Province of New Brunswick, 1989. Land Information Policy.

Shih, T.Y. and E.E. Derenyi, 1992. Orthoimage Generation in a GIS Environment. *International Archives of Photogrammetry*, Vol. 29, Part B4, p 221-223, August 2-14, International Society of Photogrammetry and Remote Sensing.

PHOTOGRAMMETRY & GPS FOR CADASTRAL LAND INFORMATION SYSTEM

by

Brig. J.S. Ahuja, Director
Pilot Map Production Plant
Survey of India

Mr. G.S.Kumar
Director (Retd.)
Survey of India

This paper attempts to review the adoption of photogrammetry for building Cadastral Land Information System based on the experiences of the authors of the paper in the various projects in India includin use of GPS in differential mode for generating digital data from ground observations for integrating with photogrammetry.

Photogrammetric method of survey is expected to cost about Rs. 100/- per hectare and will take 15 to 20 years for each of the States in India. The authors propose a phased approach in building comprehensive Cadastral Land Information System by integrating GPS technology with photogrammetry to reduce cost and time significantly, while maintaining the accuracy standards. Such a phased and integrated approach is considered appropriate in countries needing fresh/re-surveys for building Cadastral Land Information System.

INTRODUCTION

The subject of cadastral surveys is a complex one. Besides technology, it has historical, political, legal and social dimensions. In ancient India land measurement and taxation was practised. According to documentary evidence, in the first century BC, ownership of lands of grounds were numbered with descriptive information indicating cultivated, non-cultivated, plains, marshy lands, gardens, vegetables, forests, temples of Gods, irrigation works, cremation grounds, feeding houses, water availability points for travellers, places of pilgrimage, pasture grounds and roads. Boundaries of villages, forests, roads, registered gifts, charities and taxes regarding fields also were indicated in those records. During 15 - 16th centuries lands were carefully surveyed and for measurements the older units were replaced by yards and subsequently to metres.

The most important objective of the surveys was tax collection. At the end of 18th century, East India Company strengthened the surveying systems to aid revenue collection. Due to the diversity of various traditions in the country, the revenue survey system developed heterogenously. However, in recent times, area accuracies were aimed at $\pm$ 1% of the total area measured. According to the Indian Forest Act of 1863, various areas came under the Forest Revenue departments. However, the land holder had full rights of ownership subject to the payment of government revenue. With growing population and pressure on land-use the emphasis has shifted from land revenue to better land utilisation and efficient land management system. This necessitated more accurate methods of measurements in surveying.

In the present day context with pressure increasing on land, Cadastral Surveys has become a service to both the owner of the land and to the State. These Surveys will, therefore, serve multipurpose for fiscal, legal, administration and economic development. The benefits will be more of social value as there will be reduction in expenditure on solving litigation cases and may not be quantifiable in economic terms on the basis of cost-benefit ratio.

REVIEW OF THE STATUS OF CADASTRAL SURVEYS IN INDIA

Recognising the importance of this problem, Resolution No.23 was passed at the second United Nations Regional Cartographic Conference for Asia and the Far-East, held in Tokyo in 1958, recommending amongst other points that Cadastral Surveys should in all cases be based on a sound geodetic control which should be connected with the existing national geodetic net.

In India, a beginning was made in 1957 by organising regular Cadastral Survey Conferences regularly. The first Conference called the 'First All India Cadastral and Forest Surveys Conference', which was held at Delhi, on 9th March 1957 felt an immediate need for an enactment of the Standards Weights and Measures Act 1957 to be made obligatory for the Central and State Governments and other agencies to adopt metric weights and measures, which also involved changing the scale of Cadastral and Forest Survey Maps and the chains being used for survey measurements. The Second Conference renamed as 'All India Cadastral Surveys Conference' held at Mussoorie (U.P) in June 1964 has stressed on the following objectives, regarding Cadastral Surveys.

- Modernisation of techniques and equipments
- To bring about better coordination between the Survey of India and the State Survey Departments
- To consider and provide adequate training facilities for the Cadastral Survey personnel

Cadastral maps are large scale maps and need greater effort in the field in terms of equipment and personnel. They are more expensive as compared to topographical maps when prepared in a scientific manner after following surveying principles. The cost factor can be considerably reduced by cutting down the predominant cost of field work. Aerial photography has come to the rescue in achieving this aim. There is no other alternative but to use various photogrammetric methods. These methods offer superior products with a large volume of information in a short time. The present day cadastral maps should not limit themselves to the narrow definitions but should serve as land information maps required for planning at all levels. They should even form base for project surveys and even topographical surveys, where possible in due course of time.

ANDHRA PRADESH EXPERIENCE

Large scale cadastral surveys in the State of Andhra Pradesh were conducted during the last quarter of 19th Century and the first quarter of this Century. For historical reasons, different systems of surveying the lands and preparing basic land records were adopted in the two regions of the state, in Andhra and Telangana areas, which were then parts of the Madras Presidency and the Nizam State respectively.

The earliest cadastral records in the State are now more than hundred years old and have mostly outlived their utility. They are not commensurate with the present day requirement and do not reflect the aspirations of the land holders. The procedures adopted for maintenance and updation of these records are archaic, time-consuming and often result in the very failure of the maintenance system. Consequently, the outdated land records are not in a position to satisfy the growing demands of both the public as well as the Government for the following reasons -

- Gigantic workload due to lack of updating since decades.
- Limited survey potential and meagre fiscal resources of the States.
- Inordinate delay in completion of Survey/Resurvey process using conventional method.
- Pressing demands from the ryots for up-to-date records.
- Implementation of Rural Development Programmes.
- Pending legal disputes regarding ownerships for want of reliable records.
- Limited land vis-a-vis phenominal population-growth.
- Soaring land value.
- Restricted requirement of number of copies.
- Need to develop Land Information System.

ACCURACY STANDARDS

The Survey Manual of Departmental Rules issued by the Govt. of Andhra Pradesh lays down the following accuracy standards.

Measurements of diagonals and field and subdivision boundaries -

In lines of 5 chains (100 m) and under, a difference exceeding 3 links (0.6 m) is an error.

In lines of over 5 chains (100 m), a difference of more than one link (0.2 m) per chain is an error upto a limit of 10 links (2 m), but any difference of over 10 links (2 m) is an error.

Measurements on diagonal lines, where the cross staff has been used, or offset measurements.

In lines of 5 chains (100 m) and under, a difference exceeding 4 links (0.8 m) is an error.

In lines over 5 chains (100 m), a difference of more than one link (0.2 m) per chain (20 m) is an error upto a limit of 10 links (2 m), but any difference over 10 links (2 m) is an error.

The feasibility of achieving the above accuracy standards has been practically established through a pilot project recently carried out by the Pilot Map Production Plant, in a village in Rangareddi District, Hyderabad. An excerpt of the results achieved is shown in Appendix I. Despite the fact that the 1:10,000 scale photography used is more than 8 years old and certain changes have taken place on the ground during the intervening period and some of the bunds are non-existent, the results achieved are quite convincing, leading to the firm conclusion that by impressing upon the ryots to raise ridges, wherever not existing, on their boundaries, followed by fresh aerial photography could lead to achieve better results.

These accuracy standards can be conveniently achieved by employing numerical methods of stereophotogrammetry. Even precision analog instruments, coupled with automatic data recording device can be used. However, work-stations with PC- based digital systems, which are capable of processing numerical data, extracted from stereo models, will be a better configuration as this could dispense with manual drills of absolute orientation, thereby eliminating residual errors on account of sources like projection and plotting of control points, scaling error due to bias or personal judgement. The photogrammetric data could be captured either from a relatively-oriented model or from an analytically-absolutely oriented model. In the former case, the model co-ordinates could be transformed to terrain terms using sophisticated software, and in the latter case, after achieving numerical absolute orientation, the model points also could be got transformed to terrain terms using the transformation parameters. From these ground co- ordinates of all corners and turning points of the holdings, a blue-print for the village map could be prepared through computer design files, which could conveniently be used for ground verification, in case the photographic products are not preferred. From the ground verified information, the extraneous bunds, not forming boundary limits of any holdings, could be eliminated and the data in the computer could be edited and updated. Using the above and various attributes collected, the individual plot map could be plotted giving all desired dimensions, area of the plot and land-use data, a specimen copy of which has been given as Appendix II. By consolidating all these plot maps pertaining to one village, a village map and combining the constituent villages, a mandal map, on a convenient smaller scale, can be produced as computer outputs. The advantages of this modern technique will be that the data could be stored in digital form on floppies, facilitating easy periodical updating by the Cadastral Survey Organization, after acquiring necessary technical know-how through interaction with the survey department. This technique can supply the end-products in the form of soft and hard copies of computer output. Other advantages are storage-convenience, easy retrieval, editing facilities, and scope for supplementing it with multi-cadastre data for other user-organizations, as and when required.

GOA CADASTRAL SURVEYS

This project was undertaken during 1971 covering an area of 8,000 hectares consisting of 16 villages. The aerial photography was taken in January, 1970 with RC 5a camera, 11.5 cm focal length on 1:6,700 scale approximately. Enlargements on the scale of 1:3,000 were also used. Before the aerial photography, demarcation and pre-marking was carried out. In view of the practical problems involved in pre-marking every property holding, it was decided to do the pre-marking only for survey number boundaries i.e., a group of holdings. Most of these boundaries followed field bunds making identification easy. The pattern of pre-marking consisting the trenches in the shape of a cross, the legs being 50 cms x 15 cms with a gap of 75 cms square in the middle. Subsequently, model control was provided using electronic distance measuring (EDM) instruments for planimetric points and spirit-levelling for heights. Pre- marked points, clearly identifiable on the photography, were included as control points. Photogrammetric processing was carried out on the basis of photographs verified on the ground for fixing of individual land holdings. Plotting was done on wild A7 and A8 stereo plotters. Co-ordinates were read for all the pre-marked points, besides some other sharp points. This material was used for field completion. Plotting was carried out village-wise on 1:1,000 scale. Village maps on 1:5,000 scale were also produced. Record of Rights (ROR) was also generated. Out of 8,000 hectares of the area attempted 4,000 hectares consisted of built-up and covered areas involving extensive ground work. The cost per hectare by photogrammetric method worked out to Rs.64/- as against Rs.72/- by conventional method. Though this project was a technological success, the introduction of modern methods suffered set back due to other factors.

ANGUL-NALCO CADASTRAL SURVEY PROJECT IN THE STATE OF ORISSA

This is a collaborative project between Survey of India and Department of Land Records and Surveys, Govt. of Orissa, funded by Ministry of Rural Development, Government of India. The objective is to survey 400 villages. Aerial photographs on 1:10,000 scale have been used. Ground control has been provided using EDM instruments and Sprit-levelling. The preliminary results showed an accuracy of 0.5 metre which compare favourably with the accuracy requirements in the traditional system. An integrated approach by using the co-ordinates generated by stereo plotters for a village map/land parcel map and other connected maps as required by the local administration are being generated with greater speed and economy. The entire work is expected to cost about Rs.150 lakhs.

GPS FOR CADASTRAL SURVEYS

Global Positioning System (GPS), which provide coordinates and heights of points with the signals from satellites, is emerging as a powerful technology for application in the cadastral surveys also. If used in an organised manner, this can replace the present system of providing control using EDM instruments. Sample observations were made in December '94 using Motorolas Differential GPS. The distances computed from the observed coordinates compared favourably with the traditional accuracy standards. Maximum error was in the order of 0.65 metres. If GPS is well integrated with photogrammetric technology, building of Cadastral LIS will become economical and speedier while meeting the accuracy standards. One of the major advantages of GPS technology is that this can be used during day and night and in all weather conditions. It also requires very short programmes of familiarization for using the existing manpower.

CONCLUDING REMARKS

Planning Commission as part of its planning approach, considers spatial planning as related to the choice of locations in a region from where wide range of socio economic infrastructure and surveys facilities are to be provided effectively. Having appreciated importance of spatial dimension of planning, Govt. of India is very keen on generating Land Information System through digital technology. On a national basis, this involves huge investments of money, manpower, effort and time. Department of Rural Development, Govt. of India is attempting to undertake a number of pilot projects in different parts of the country. In a couple of years, it is expected that results will be available for evolving standards for digital photogrammetric technology for cadastral surveys in India.

REFERENCES

Dale P.F. Cadastral Surveys within the Commonwealth; His Majesty's Printing Office.

"Fiftieth Anniversary Highlights" Map Accuracy Specifications. Photogrammetric Engineering and Remote Sensing, Feb. 1984:237 - 240.

Kumar G S, et.al 1985 rectified photographs for cadastral surveys - a case study, of India Cadastral Survey Seminar, Dehra Dun.

Nagarajan B (1988) Cadastral and Land Information System Technical Publication No.8842 R&D Directorate, Survey of India.

Paul E Norman, 1965. Photogrammetry and the Cadastral Survey. A33 ITC Publication : 65 pages.

Pillai K N G K and Ahuja J S, Cadastral Surveys using Digital Photogrammetric Workstations present at 13th INCA Congress.

Quinn A O, 1983. Legal Aspects of Photogrammetric Measurements for Surveying and Mapping. Paper published at New York State Association of Professional Land Surveyors Conference. Photogrammetric Engineering and Remote Sensing, 50(4):453-456.

Robert Burtch and Khagendra Thapa, 1990. Multipurpose Cadastral Accuracy Standards. Surveying and Land formation Systems, 50 (4) : 271 - 277.

Wander Weele A J, 1974. Photogrammetry for Cadastral Surveys. Paper Presented to the United Nations' Inter regional Seminar on Cadastral Surveying and Urban Mapping. The ITC - Journal 1974 - 5.617 - 627.

Zhilin Li Hill, C.T., AZZIZI, A. and CLARK , M.J 1992. Exploiting the Potential Benefits of Digital Photogrammetry Some Practical Examples. Paper read at the Thompson Symposium. Photogrammetric Record, 14(81) : 469 - 475.

COMPARISON OF MEASUREMENTS
(metres)

Appendix I

Survey No.	N		S		W		E		REMARKS
	Ground	Photo	Ground	Photo	Ground	Photo	Ground	Photo	
1	78.4	75.5	68.4	67.8	9.0 14.4 20.4	9.0 12.6 22.6	56.0	57.4	North-eastern corner point is not well-identifiable.
2.	68.4	67.7	53.8	54.4	15.6 29.6	19.6 24.2	49.0 4.8	49.1 3.9	Total length of western side agrees
3.	38.2	38.0	24.8	27.0	49.0 56.0	49.2 57.4	110.4	109.8	Total length; intermediate positions not correctly identifiable.
4.	37.2	36.3	8.2 21.4	8.0 21.2	39.8	39.5	34.8 7.6	-- --	Eastern points could not be correctly identified.
5.	28.0 18.6	27.0 19.0	37.2 12.2	36.6 14.6	29.0	--	5.6 8.0 24.2 12.2	5.9 8.7 22.4 13.9	-do- Total length of Eastern side agrees
6.	24.6	24.8	28.0	27.0	41.6	--	40.8	41.1	Western endpoints could not be located correctly.
7.	17.4	17.3	18.6	19.2	40.8	--	33.0 14.2	--	SW corner could not be located correctly. Intermediate point on the Eastern side could not be lo cated correctly.
31	14.4 3.8	14.2 4.0	13.8	14.3	54.0	54.2	33.8 29.4	34.0 29.9	-- --
32	15.6 14.4	16.2 13.2	25.2 3.8	25.2 4.8	29.4	29.9	29.6	29.6	--
47	25.2	25.0	20.2 5.6	18.9 6.0	33.8	33.5	33.8	32.7	--
48	25.6 13.8 20.2	-- 14.3 18.5	23.6 25.0	-- 27.0	29.0	28.9	18.0	18.1	Positions of Western endpoints are different on photo from ground.

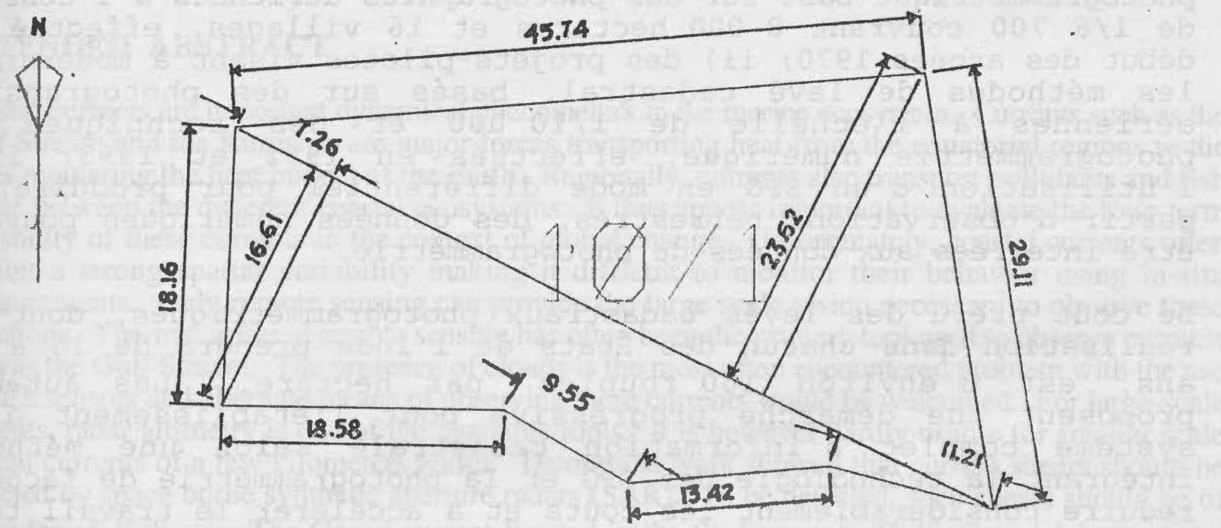
PLOTMAP

REV. SUR. No. : 10/1 FIELD INDEX NO. 17

OWNERSHIP : Krishna Reddy ADDRESS : H-No 45/6

VILLAGE : Manikonda

MANDAL : Rajendranagar DISTRICT : Rangareddi



SCALE 1:500 AREA = 0.11 Hect.

CO-ORDINATES (In metres)				LAND-USE INFORMATION
POINT	X	Y	Z	
1	2571167.9	738693.3		Soil - Clay
2	2571213.5	738697.1		Crop - Paddy
3	2571218.1	738668.4		Irrigation - Well
4	2571207.3	738671.2		Electrification - No
5	2571193.9	738670.1		Revenue - Rs.8.00
6	2571185.6	738674.9		
7	2571167.1	738675.2		

OFFICE SEAL SIGNATURE

Établissement d'un système d'information cadastrale au moyen de méthodes photogrammétriques et d'un SPG

Résumé

Le présent exposé traite de l'adoption de procédés photogrammétriques pour l'établissement d'un système d'information cadastrale, procédés dont se sont servis les auteurs dans le cadre des projets suivants en Inde : i) un levé cadastral photogrammétrique basé sur des photographies aériennes à l'échelle de 1/6 700 couvrant 8 000 hectares et 16 villages, effectué au début des années 1970; ii) des projets-pilotes visant à moderniser les méthodes de levé cadastral, basés sur des photographies aériennes à l'échelle de 1/10 000 et des techniques de photogrammétrie numérique, effectués en 1992 et 1993; iii) l'utilisation d'un SPG en mode différentiel pour produire, à partir d'observations terrestres, des données numériques pouvant être intégrées aux données de photogrammétrie.

Le coût prévu des levés cadastraux photogrammétriques, dont la réalisation dans chacun des États de l'Inde prendra de 15 à 20 ans, est d'environ 100 roupies* par hectare. Les auteurs proposent une démarche progressive pour l'établissement d'un système complet d'information cadastrale selon une méthode intégrant la technologie des SPG et la photogrammétrie de façon à réduire considérablement les coûts et à accélérer le travail tout en respectant les normes de précision. Cette démarche facilite également l'utilisation des résultats obtenus depuis le tout début de la mise en oeuvre du projet. Cette démarche progressive et intégrée convient particulièrement bien aux pays où l'établissement d'un système d'information cadastrale nécessite de nouveaux levés.

* Un dollar américain vaut environ 31,40 roupies indiennes.

OBSERVATIONS OF A COASTAL CURRENT USING ERS-1 SAR

PIERRE LAROUCHE
Institut Maurice-Lamontagne, Ministère des Pêches et des Océans
B.P. 1000, Mont-Joli, Québec
Canada G5H 3Z4

Commission II, Group 4

KEY WORDS: SAR, ERS-1, Currents, Circulation, Oceanography

EXTENDED ABSTRACT

Coastal currents are important dynamical phenomena's in the marine ecosystem. Currents such as the Gulf Stream and the Kurushio are major forces transporting heat from the equatorial regions to the poles regulating the heat budget of the Earth. Regionally, currents also transport pollutants and fish larvae between the different coastal ecosystems. It thus appear important to evaluate the long-term variability of these currents in the context of global change. Unfortunately, coastal currents often exhibit a strong spatial variability making it difficult to monitor their behavior using in-situ measurements. Only remote sensing can provide the large scale vision necessary to observe these variations. Thermal infrared remote sensing has often been the primary tool used to observe currents such as the Gulf Stream. The presence of clouds is the most often encountered problem with the use of such sensors and alternate means of observing these currents would be welcomed. For large-scale currents, radar altimetry is one of the available tools. It is however hardly usable for smaller scale coastal currents of a few kilometers width. Theoretical work showed that current shears should be detected by space borne synthetic aperture radars (SAR). To be detected, such shears should be of the order of 10^{-4} s^{-1}. The Gaspé current located in the St. Lawrence Gulf (Canada) is a coastal current that meets this criteria. The goal of this project was thus to verify the possibility for a space borne instrument to detect the location of coastal currents by comparing active microwave C-band SAR data with concurrent in-situ measurements of current shears.

Two ERS-1 MLD SAR images (September 25 1991 and May 19 1992) were used for the first phase of the project. In-situ measurements consisted of a series of transects perpendicular to the mean current direction made during the two satellite overpass using an acoustic Doppler current profiler (ADCP). Salinity and temperature profiles were also gathered during the transects to determine the density structure of the current.

Image processing first consisted of locating the transects location and extracting sub-images corresponding to these areas. An adaptive Frost filter was then applied on these sub-images to remove speckle while preserving the high frequency features present. Grey level values were then extracted from the images along a line corresponding to the ship track and smoothed using a running mean of size 10.

The analysis of the data from the three transects indicates that the ERS-1 SAR detected the location of the Gaspé current boundary when environmental conditions were favorable (winds between 1 and 5 m s^{-1}) thus opening the possibility of using this tool to monitor coastal currents. It was thus shown that shears of 1 x 10^{-4} s^{-1} were indeed detectable as predicted by theory. The observations in this experiment were however made only with a current flowing roughly in the SAR range direction. Further observations should thus be made with a broader set of currents (including azimuth traveling currents) to further evaluate the capacity of a space borne SAR to detect this type of dynamical features.

OBSERVATIONS D'UN COURANT CÔTIER AU MOYEN DU RADAR À SYNTHÈSE D'OUVERTURE DU ERS-1

OVERVIEW OF THE WORK ON THE CLASSIFICATION OF SAR SHIP IMAGERY PERFORMED AT DREO

Robert Klepko
Defence Scientist
Defence Research Establishment Ottawa
3701 Carling Avenue
Ottawa, Ontario, K1A 0Z4

KEYWORDS: Synthetic Aperture Radar, SAR Images, Ships, Classification, Computer-based tools, Operator-machine interface

ABSTRACT

The Airborne Radar and Navigation Section at the Defence Research Establishment Ottawa (DREO) has developed an airborne Synthetic Aperture Radar (SAR) that is capable of generating, in real-time, high-resolution radar images of ocean-going ships. The acquired SAR images of ships must be classified as to the type of ship, and this is performed by a Radar Operator. Classification must be done in real-time and with precision. Given that there are an abundance of images, and that these images can be generated at a high rate, it is difficult for a Radar Operator to maintain an accurate classification without the use of some aids.
A set of Operator aids, in the form of a computer-based graphics system, is currently being researched and developed. This system is called the Radar Operator Interactive Classification and Training System (ROICATS). ROICATS is composed of three main components: computer-based tools, operator-machine interface features, and automated target classification algorithms. This paper gives an overview of the work done on each of these three components.

RÉSUMÉ

La section du radar aéroporté et de la navigation du centre de recherches pour la défense à Ottawa a mis au point un Radar à ouverture synthétique (ROS) qui génère des images radars à haute résolution de navires de haute en temps réel. Un opérateur de radar doit classer les images ROS des vaisseaux en catégories. La classification doit être faite en temps réel et avec précision. À cause de l'abondance d'images et parce qu'elles sont générées tres vite, l'opérateur du radar peut difficilement les classer correctement sans aide.
Nous mettons au point un ensemble d'aides à l'opérateur, sous la forme d'un système informatique graphique. Ce système est appelé le Système d'entraînement et de classification interactive avec l'opérateur radar (SECIOR). Le SECIOR comprend trois composantes: outils informatiques, interface usager et algorithme de classification automatique de cibles. Ce rapport résume le travail accompli pour chacune des trois composantes.

1.0 INTRODUCTION

The Airborne Radar Section at DREO has developed an airborne SAR system, called SPOTlight SAR, which is capable of generating, in real-time, high-resolution radar images of land targets, and ships at sea. This SAR system is actually a new operational mode that will be incorporated into the AN/APS-506 Maritime Search radar on the CP-140. The AN/APS-506 radar, as it exists now, has three conventional modes of operation which are used for: periscope detection; weather watch and navigation; and long range detection of ships. It is possible to track a single target in all three modes. Because of the conventional search radar's coarse spatial resolution, it is not possible to accomplish the main requirement for classification, which is to determine whether a ship is combatant or noncombatant. Hence, the SAR mode was introduced to provide both this capability, and a capability for high-resolution land imaging.

The SAR mode can be sub-divided into one of three sub-modes. In the first sub-mode, called the SPOTlight mode, the antenna is maintained fixed on a specific target, a certain amount of radar returns are accumulated, and a very high resolution radar image is generated from this data. With this image, detected targets can be classified. In an adaptive version of this mode, moving ship targets on the open ocean are tracked and compensated for their motion. In nonadaptive SPOTlight mode, a high-resolution image of a small area of land is generated. This image can be analyzed for the presence of man-made targets, and is particularly useful for harbour surveillance.

The second sub-mode is called Range-Doppler Profiling (RDP). In this mode the antenna is also maintained fixed on a specific target, but instead of a single radar image being generated, as in the SPOT mode, a continuous series of high resolution radar images of ships are generated in real-time. Because of the constraints of real-time processing, this mode does less adaptive processing than the SPOT mode, but the Radar Operator is provided with a greater variety of images with time. This mode is not used for land imaging. In both RDP and adaptive SPOT modes, either stationary or moving ships can be imaged.

In the third sub-mode, which is called the STRIPMAP mode, a continuous strip, or map, of high resolution radar imagery is generated. In this mode, spurious motion of targets is not compensated, so it is not used for imaging ships on the open ocean, where the motion of the ship would severely degrade the image. Its chief application is for land imaging.

Because STRIPMAP imagery is generated in real-time, a Radar Operator must view a large amount of data at a high rate while trying to detect man-made targets in the images. If a man-made target is detected, it must be analyzed for various attributes, such as its shape and location. This analysis can be performed either immediately, while still viewing the STRIPMAP imagery for the presence of other targets, or more likely, by saving an image frame and recalling the image for analysis at some future time. Analyzing these images can be very demanding on an Operator, especially after working for any great length of time. To alleviate the potential fatigue that an Operator may encounter, some form of aid should be provided to the Operator during the analysis process. The aid should consist of automated target detection and localization procedures, and map files and overlays.

The remainder of this paper discusses work done in the classification of RDP and (adaptive) SPOTlight images of ships.

The appearance of SPOT or RDP images of ships is a function of many variables, such as: the location and type of scatterers on the target, the SAR system frequency, the radar-to-ship viewing angle, the amount and type of sea-induced ship motion, and the cross-range resolution. The appearance of the images of stationary land targets also varies with the location and type of scatterer, frequency, aspect, and cross-range resolution. However, they do not experience the large changes in the orientation of the image plane which is seen with moving ships. Because of the dependence of the appearance of a SAR image (i.e. SPOT or RDP) on the variables mentioned above, the number of possible images that can be generated from any one target is large. Since the potential number of targets is also large, the total number of possible SAR images is very large indeed. It is the job of the Radar Operator to understand the abundance of SAR images that are generated, and attempt to make a speedy and correct decision as to either the type or class of target. Since it would be very difficult for an Operator to remember all possible images that might be encountered, some form of aid should be provided for the classification process.

The Airborne Radar Section at DREO is currently developing a computer-based graphics system to aid the Radar Operator in performing speedy and accurate classification. This system is called the Radar Operator Interactive Classification And Training System (ROICATS). The present system is composed of three main components: a graphical-user-interface which supports the use of numerous computer-based tools, operator-machine interface features, and automated target classification algorithms. The training aspect of the system has yet to be developed.

The tools allow image features to be extracted and various forms of information to be manipulated. Typical tools include an electronic ruler for measuring the distance between image features, an extraction tool to extract a subset of an image, and image display and manipulation options. Information, such as the physical dimensions of a ship, is stored in an interactive database.

The operator-machine interface provides the structure for simplified interaction between the Radar Operator and ROICATS. For example, most of the system interactions can be done through the use of just a mouse or trackball, while avoiding the tedious use of a keyboard.

The automated target classification algorithms analyze the SPOT and RDP ship images to determine a list of ship classes that may match the target in the image. The list should be much smaller than the complete list of typical classes, so that the Operator only needs to consider a subset of the possible classes. This speeds up the classification process and improves the accuracy since there is less of a chance of matching the target to an incorrect ship class.

ROICATS is being developed on a SUN SPARCstation 370GX engineering-graphics workstation. The computer-based tools and operator-machine interface features are realized using the PHIGS (Programmer's Hierarchical Interactive Graphics System) and C programming languages, while the automated algorithms are being realized in FORTRAN 77. The automated algorithms are currently being developed on a 486-based PC, but will eventually be implemented on the SPARCstation.

The next section of this report discusses the types of automated classification algorithms that have been developed for eventual use within ROICATS. Section 3 discusses the types of computer-based tools that are being considered. Section 4 discusses the operator-machine interface features that are being implemented. The paper concludes with a discussion of the future work required to complete the development of ROICATS.

2.0 AUTOMATED SHIP CLASSIFICATION ALGORITHMS

Over the past eight years, the Airborne Radar section has been researching and developing numerous algorithms that can be used to automatically classify high-resolution radar images of ships. Only a brief description of the most noteworthy of these algorithms is given.

The various classification algorithms can be divided into two general catergories. The first category is based upon a robust technique that uses coarse image features to make a classification. Various coarse features are extracted from an input ship SAR image and are compared to a ship library containing the same features. There are basically six types of coarse features extracted: the number, level and location of major peaks in the image; the relative range location and height, in cross-range, of any vertical structures evident in the image; the estimate of the minimum ship length; the presence of the bow of the ship; the outlines of the sides of the ship; and the presence of a rotating antenna. After extracting values for these features, an Expert System (ES) methodology (i.e. a set of rules, a knowledge base, and a rule control strategy) is used to guide the comparison of features. Since the features are coarse and not necessarily persistent, there is a good chance that several ship classes will be declared to be valid matches by the ES. It is required that the ES output a list of candidate ships which is much shorter than the initial list of all possible ship classes. The list of candidate ship classes is presented to the Radar Operator and fed into the second category of classification algorithms. By having to consider a much smaller subset of the library of ships, the subsequent classification process should be faster and there should be less chance for error. Incorporated into the ES is a techique which automatically determines the quality of an input SAR image. This technique basically allows the ES to filter out those images which are considered too poor for classification.

The second category of algorithms base their decisions on the actual pixel values in the input SAR image. There are three different techniques being developed in this category: a combinatorial optimization technique (sometimes referred to as simulated annealing), a decision theoretic technique, and neural networks.

The combinatorial optimization technique takes a computer-based block model of a ship (an example of such a model is shown in Figure 1), and simulates its SAR image for various values of the following four parameters: the three angular velocities, roll, pitch, and yaw; and the ship heading with respect to the SAR line-of-sight. The simulated image is compared to the real input ship image using a two-dimensional correlation, and if the comparison is poor, the value of one of the parameters is modified and another image is simulated. The schedule, or method of selecting a parameter and its new value, is the heart of the combinatorial optimization technique. This technique finds the optimum set of parameter values so that the best match between the input and simulated images can be obtained.

The decision theoretic method uses templates that represent incrementally learned simulated SAR images of ships. For each class of ship there are several templates, each representing a range of aspect angles corresponding to 45 degrees. Simulated SAR images of a ship are used to incrementally learn a template, also called a discriminant vector, that is representative of that ship for the range of aspect angles in question. Only after each vector is created can the decision theoretic method perform a classification. A classification decision is based upon computing the two-dimensional correlation between the input ship image and each discriminant vector. The vector which gives the highest correlation is considered to represent the class of the ship in the input SAR image.

Neural Networks are also being evaluated as part of the second category of algorithms. These networks promise to provide quick recall of learned pattern associations with high recognition performance, even when noisy or incomplete input patterns are encountered. Research and development is being performed primarily with the BackPropagation Network and derivatives of it.

Figure 2 shows a block diagram of the overall classification system, ROICATS. The operation of ROICATS is as follows. A real SPOT or RDP image of a ship is fed to the Expert System which then uses its library of ships' features to determine a list of potential or candidate ship classes that may represent the input image. This list is fed to each of the three detailed analysis algorithms, as well as to the Operator-Machine Interface. This interface contains the computer-based tools and operator-machine interface features which are discussed in the following two sections. While the three algorithms perform their respective automated operations, the Radar Operator performs a manual classification.

Since none of the three classification algorithms can be expected to achieve 100 percent recognition performance, their outputs may not be correct and/or the same. If, however, their independent outputs could be combined, in a manner which reflects the probabilities of the different possible ship classes, then enhanced classification could be achieved. However, since these probabilities are not known, no assumption of the initial probability distributions of these classes can be made. The Dempster-Shafer method of evidential reasoning allows probability estimates to be made using the independent output values, without knowledge of their initial distributions (Blackman, 1986). The Dempster-Shafer method is considered an adequate means of combining the output of the three classification algorithms, along with tactical and intelligence information, to obtain the decision on the final output ship class. This ship class is sent to the Radar Operator via the Operator-Machine Interface.

3.0 COMPUTER-BASED TOOLS

Computer-based tools provide the Radar Operator with the means to extract, or relay, various pieces of information from, or to, ROICATS. The extracted information aids the Operator in performing the classification analysis, while the relayed information allows certain operations to be performed by ROICATS. Generally speaking, the tools available to the Operator should only provide information that is essential to the classification of the input SAR image. Also, only the number of tools required for a reasonable degree of system flexibility should be provided since too many tools may be confusing to the Operator, particularly in a real-time scenario. To date, eight groups of tools have been devised, and these are listed in Table 1. The purpose of each tool group is as follows.

The SAR images output by the SAR signal processor will contain the ship radar returns of interest as well as backscatter from the sea, and perhaps other ships or man-made objects. All image data other than the ship radar returns may be regarded as noise since they are not informative to the classification process. Thus, it would be advantageous to the classification process to be able to extract only the region in a SAR image corresponding to the radar returns of the ship. Two schemes are available to the Operator as indicated in Table 1.

After the portion of the SAR image corresponding to the radar returns of the ship has been extracted, the Operator may wish to enhance the appearance of the image by using one of the image processing algorithms listed in Table 1.

After enhancing an image, it can be displayed on a computer monitor using a set of display tools which provides various display styles. These styles are listed in Table 1.

To help access and manage the library of ships, a set of Database Management tools must be provided. These tools allow the Operator to retrieve any piece of information that is stored for any ship, or to add information about any ship to the library.

The displayed image can be manipulated using any one of the tools listed in Table 1. Also, the computer-based models of ships can be manipulated by rotating, scaling or translating them on the computer screen. These displayable models are used to help the Operator to better visualize the true appearance of the various ship classes.

Once the SAR image is displayed in an ideal format, various distinguishing features, used for classification, can be extracted. To extract various features from the displayed ship images, a set of Image Feature Extraction tools are provided. A tool to measure the absolute distance, in range and cross-range, between two Operator-selected points, is very useful. An estimate of the minimum length of a ship, based upon Operator inputs, is another useful tool.

Table 1 lists two other tools which can aid in the classification process. They are an Overlay tool to allow the Operator to directly compare real ship images with computer-based models of the ship's superstructure, and a Simulator tool to allow the Operator to simulate a SAR image of a particular ship to determine if that ship may be the one in the image.

4.0 OPERATOR-MACHINE INTERFACE

The Operator-Machine Interface (OMI) provides features which make ROICATS user friendly and allows information (graphical and alphanumeric) to be presented to the Operator in an effective and efficient manner. These features maximize Radar Operator performance and minimize conditions that could lead to operator error. Table 2 lists the types of features which ROICATS incorporates.

The OMI feature which is first and foremost is the ability to simultaneously display various types of information on a computer screen. Figure 3 illustrates the various computer screen display options available to the Operator. All of the display options are composed of four basic windows: the Static Menu, the Dynamic Menu, the Softkeys Menu, and the Display Area.

The first window, called the Static Menu, is used to display alphanumeric information relevant to the images currently displayed. This information includes: the image file name, the aspect and depression viewing angles of the displayed image, the scale factor, the colour mode, and the display style mode.

The second window, called the Dynamic Menu, is used to display all the alphanumeric information output by a computer-based tool (e.g. database contents and HELP descriptions). Another form of information that can be displayed in this window is the output from the automated ship classification algorithms (e.g. a list of candidate ship classes determined by the Expert System, the ship class output from each of the three automated algorithms, and the combined output from Dempster-Shafer algorithm (see Figure 2)).

The third window, called the Softkeys Menu, is used to display all of the computer-based tools that the Operator can choose from. The tools are represented by software programmable buttons on the computer screen. A button is selected by moving the cursor of a mouse or trackball over the button and pressing the enter or select key on the mouse or trackball.

The fourth window, called the Display Area, is used to display SPOT, RDP or STRIPMAP images, computer-based ship models, digitized ship photos or digitized maps. While the first three windows are fixed in size, the fourth window can be divided into one to four subwindows or areas, each capable of displaying an image. With a computer display screen size of 1280 (wide) by 1024 (high) pixels, the Static Menu should be about 256 by 150 pixels, the Dynamic Menu should be about 256 by 724 pixels, the Softkeys Menu should be about 256 by 150 pixels, and the Display Menu should be about 1024 by 1024 pixels.

The remaining OMI features listed in Table 2 are self-explanatory and will not be repeated here. Suffice it to say that this list of features represents a minimum set of mandatory features required for ROICATS.

5.0 CONCLUDING REMARKS

This paper has presented an overview of the work on the classification of SAR images of ships performed at DREO. This work has resulted in the development of a Radar Operator Interactive Classification And Training System (ROICATS). The three main components of ROICATS were described. These components are: computer-based tools, operator-machine interface (OMI) features, and automated target classification algorithms. Four different classification algorithms were briefly discussed, eight tool groups were briefly described, and five OMI features were presented.

The present development status of ROICATS can be summarized as follows. All OMI features, except for Online Guidance, have been implemented. The automated classification algorithms are at the following stages of completion: the attributes of the Expert System have been defined, but must still be implemented and tested; the Neural Network and Decision Theoretic methods are still being tested; and the Combinatorial Optimization classifier requires some minor changes and then further testing. The Image Processing, Database Manager and Overlay tools have not yet been developed.

Anticipated future tasks, required to complete ROICATS, include: complete the implementation of all the computer-based tools, implement the Online Guidance feature, and finish testing all the automated algorithms. One other major component which will eventually be part of ROICATS is a Computer-Based Training subsystem. This subsystem will be used to familiarize Radar Operators with the methodology of classifying radar images of ships. It is hoped that the first three main components of ROICATS will be completed within one year, and the training subsystem within two years.

ACKNOWLEDGEMENTS

This work was supported by the Research and Development Branch of the Department of National Defence. The review and editing of this paper by Anthony Damini is very much appreciated.

REFERENCES

[1] Blackman, S.S., 1986. **Multiple-Radar Tracking with Radar Applications**, Artech House.

[2] Rey, M.T. and Vant, M.R., 1987. "Image Enhancement and Display Techniques Applied to SAR580 Images of Ships," Communications Research Centre Report No. 1421, Ottawa, April 1987.

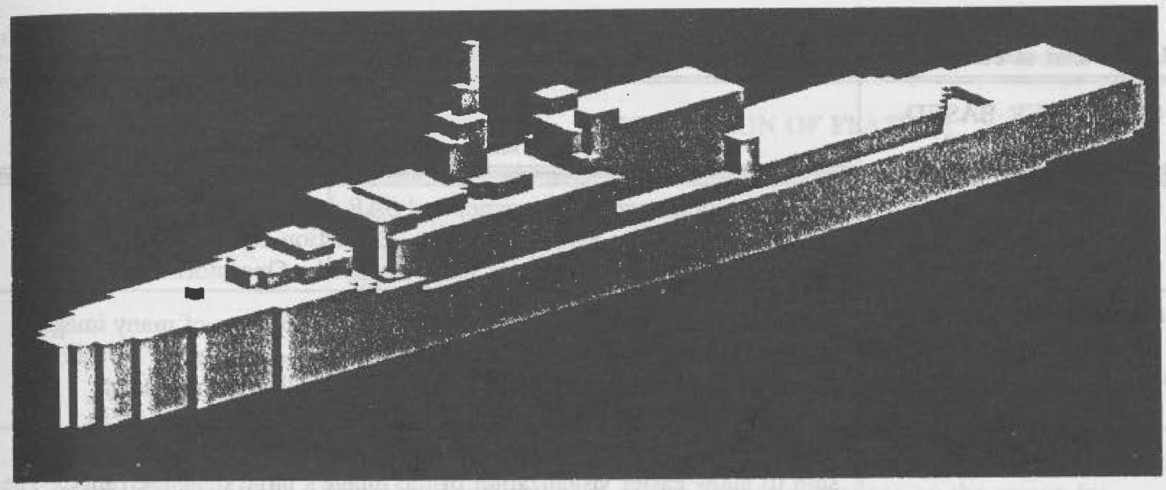


Figure 1: **Example computer-based block model representing the superstructure of a ship as seen above the waterline.**

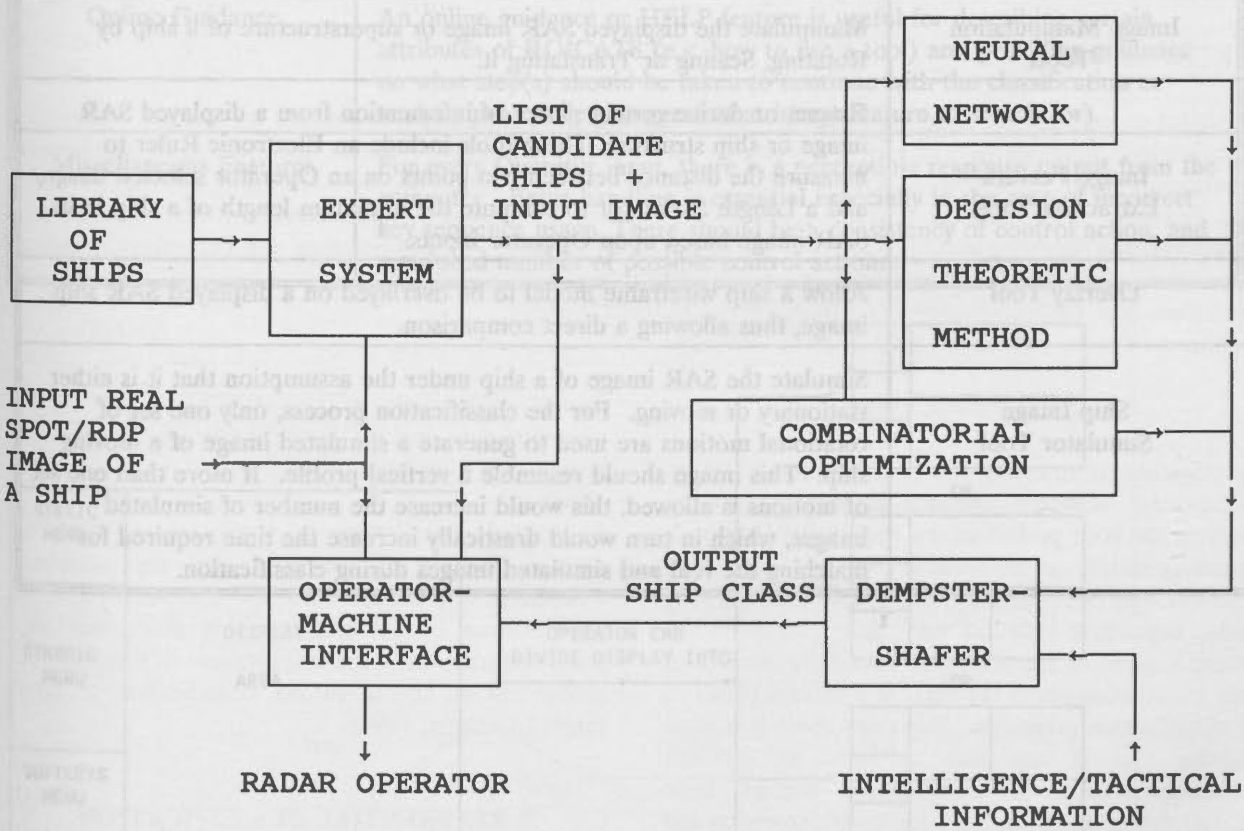


Figure 2: **Block diagram of the various major components of ROICATS.**

TABLE 1: List of computer-based tool groups to be used in ROICATS.

COMPUTER-BASED TOOL GROUP	DESCRIPTION OF TOOL(S)
Ship Image Extraction Tools	Extract a sub-image from a generated SAR image. The sub-image represents the radar returns of a ship. Extraction can be done either automatically by ROICATS, or manually by the Operator.
Image Processing Tools	Enhance the appearance of a SAR image by using one of many image processing tools. Possible image enhancement tools include: Threshold, Laplacian, Maximum Gradient, Histogram equalization and High Frequency Contrast (see Rey et al. 1987 for a description of these techniques). These techniques are usually applied to the SAR image of a ship to allow easier visualization of the image's large dynamic range. The Laplacian technique seems to provide the best image enhancement.
Display Tools	Display a SAR image using a variety of display styles. The display styles include: 3D plots (colour or monotone), Contour Plots (colour or monotone), Range Profiles, and Surface Intensity Plots.
Database Manager Tools	Allow information on SPOT or RDP stored images, or ship structures, to be added, read, deleted and/or modified.
Image Manipulation Tools	Manipulate the displayed SAR image or superstructure of a ship by Rotating, Scaling or Translating it.
Image Feature Extraction Tools	Extract or derive certain pieces of information from a displayed SAR image or ship structure. These tools include an Electronic Ruler to measure the distance between two points on an Operator selected image, and a Length Estimator to compute the minimum length of a ship in its SAR image based upon Operator inputs.
Overlay Tool	Allow a ship wireframe model to be overlayed on a displayed SAR ship image, thus allowing a direct comparison.
Ship Image Simulator Tool	Simulate the SAR image of a ship under the assumption that it is either stationary or moving. For the classification process, only one set of rotational motions are used to generate a simulated image of a moving ship. This image should resemble a vertical profile. If more than one set of motions is allowed, this would increase the number of simulated images, which in turn would drastically increase the time required for matching the real and simulated images during classification.

Table 2: Summary of the Operator-Machine Interface Features incorporated into ROICATS

OPERATOR-MACHINE INTERFACE FEATURE	DESCRIPTION OF FEATURE
Multi-Windowed Computer Screen	The computer monitor screen consists of separate display and message areas or windows. The display area allows one or more SAR images, ship structures and digitized photos to be viewed at one time. The message area provides a listing of the available tools, system status data, and information on the displayed images. The presence of multiple windows increases the readily available information and thereby improves image classification and analysis.
Mouse/Trackball Interaction Device	A computer screen cursor can be moved around using either a mouse or trackball and placed at certain strategic locations on the screen for selecting tools, and accessing or entering information. This alleviates the Operator from the tedious job of typing all information.
Multi-Colour Display	Use of more than one colour to display the images and alphanumeric information can help to highlight or emphasize to the Operator various pieces of information. However, the number of colours is limited so as not to make the computer screen appear too busy or confusing.
Online Guidance	An online guidance or HELP feature is useful for describing certain attributes of ROICATS (e.g. how to use a tool) and providing guidance on what step(s) should be taken to continue with the classification or analysis process (e.g. suggest what image feature to search for).
Miscellaneous Features	For every Operator input, there is a perceptible response output from the computer. Error handling is essential especially in the case of incorrect key sequence usage. There should be a consistency of control action, and a reduced number of possible control actions.

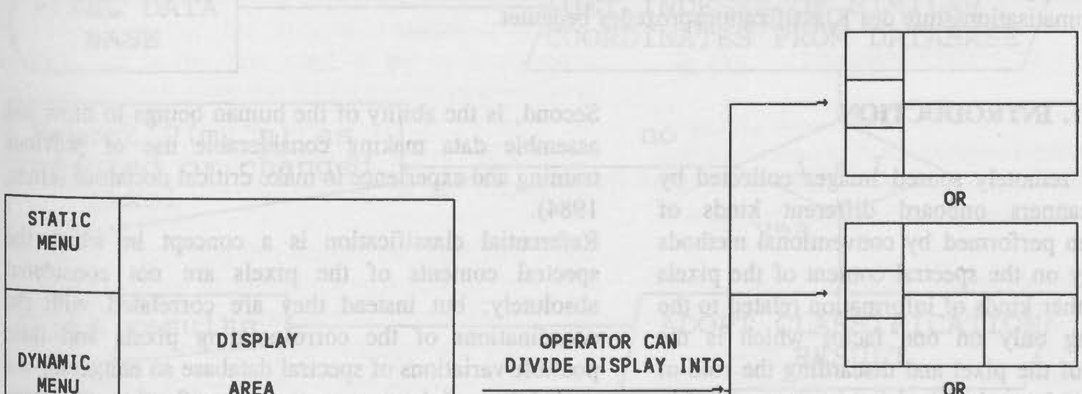


Figure 3 Various computer screen display options.

REFERENTIAL CLASSIFICATION - AN INTELLIGENCE BASED ALGORITHM

Othman Alhusain

Research Fellow, Department of Photogrammetry, Technical University of Budapest
Muegyetem rkp.3.I.24, H-1111 Budapest, Hungary

ISPRS Commission II, Working Group 2.

KEY WORDS: Image, Classification, Intelligence, Algorithm,

ABSTRACT:

This paper describes a recently developed method which can be used in the classification of pixels in satellite images. The new method named "Referential Classification" is a concept in which the spectral contents of the pixels are not considered absolutely; but instead they are correlated with a kind of ancillary data that involves the coordinations of a corresponding pixels and their possible variations of spectral data base. The referential classification process is carried out in two subsequent stages. The Application of this method in classifying remotely-sensed images resulted in higher accuracy and less demand for the user intervention during the classification course which imply higher degree of automating the classification process without sacrificing its accuracy.

KURZFASSUNG:

Im Artikel es geht um die neueste Methode, was im Satellitenbildklassifikation anwendbar ist. Die als "Referential Classification" genannte Methode ist ein Konzept, bei welchem der spektrale Inhalt der Pixel nicht absolutweise berücksichtigt wird, sondern statt dessen er mit verschiedenen Daten korreliert ist, die Koordinaten der entsprechenden Pixel und ihre Variationen des spektralen Datenbasis enthält. Der Referential Classification-Prozeß wird in zwei darauffolgenden Schritten durchgeführt. Die Anwendung der Methode in Klassifizierung von Fernerkundungsdaten ergibt höhere Genauigkeit und braucht wenigeren Eingriff während der Klassifizierung, was ohne Genauigkeitsverlust eine höhere Automatisationsstufe des Klassifikationsprozeßes bedeutet.

1. INTRODUCTION

Classification of remotely sensed images collected by multispectral scanners onboard different kinds of satellites has been performed by conventional methods depending mainly on the spectral content of the pixels while ignoring other kinds of information related to the pixels, depending only on one factor which is the spectral content of the pixel and discarding the role of other information about the pixel is a serious pitfall in the classification procedure. However, there is a clear evidence that if one like to have an efficient classification then the classification concept must simulate the human being high ability of perception and analyzing (Baxes, 1985).

The role of the human beings in the interpretation of image data is of unique and great value for two reasons: First, is the human being's ability to efficiently acquire and maintain an awareness of a universal nature; an awareness not only of the complete remote sensing data but also of the associated reference data and the relationship between the reference and spectral data. Second, is the ability of the human beings to draw and assemble data making considerable use of previous training and experience to make critical decisions (Hunt, 1984).

Referential classification is a concept in which the spectral contents of the pixels are not considered absolutely; but instead they are correlated with the coordinations of the corresponding pixels and their possible variations of spectral database so enhancing the probability of having correct classification of image pixels (Alhusain, 1992).

2. REFERENTIAL CLASSIFICATION

The algorithm implemented within this method first proceeds towards assigning Cartesian coordinates *(x,y)* to each pixel, figure 1 shows the flowchart of the referential classification method. This might be done locally where the coordinates of each pixel are valid only within the image where the assignation process has been defined, or globally where the coordinates of each pixel are related to national or even international grid coordinations. The pixel data base is used for archiving all previously measured and/or expected information about each pixel along with its coordinates within the image. The most significant factor in the pixel data base design is its ability to include multitemporal information about pixels and indices to their trends of possible change in the future because any analysis process done on pixels depends not only on the present measurements but also on measurements from the past. Coordinate assignation should be in compliance with the resolution of the imaging sensor of the satellite system; that means an image of 30 m resolution should be based on a coordinate grid of 30 m or smaller steps but not larger. The referential classification process is carried out in two subsequent stages. In the first stage a supervised classification algorithm is implemented according to the maximum likelihood rule which is maximum likelihood rule which is presented and discussed in details in image processing literature (Swain, 1978; Richards, 1986).

In this algorithm, suppose that there is a training data of an image is available for each groundcover type. This data can be used to estimate a probability $p(x/w_i)$ distribution for a cover type that describes the chance of finding a pixel belonging to a class w_i at a specific position x. This could expressed by the term $p(x/w_i)$.

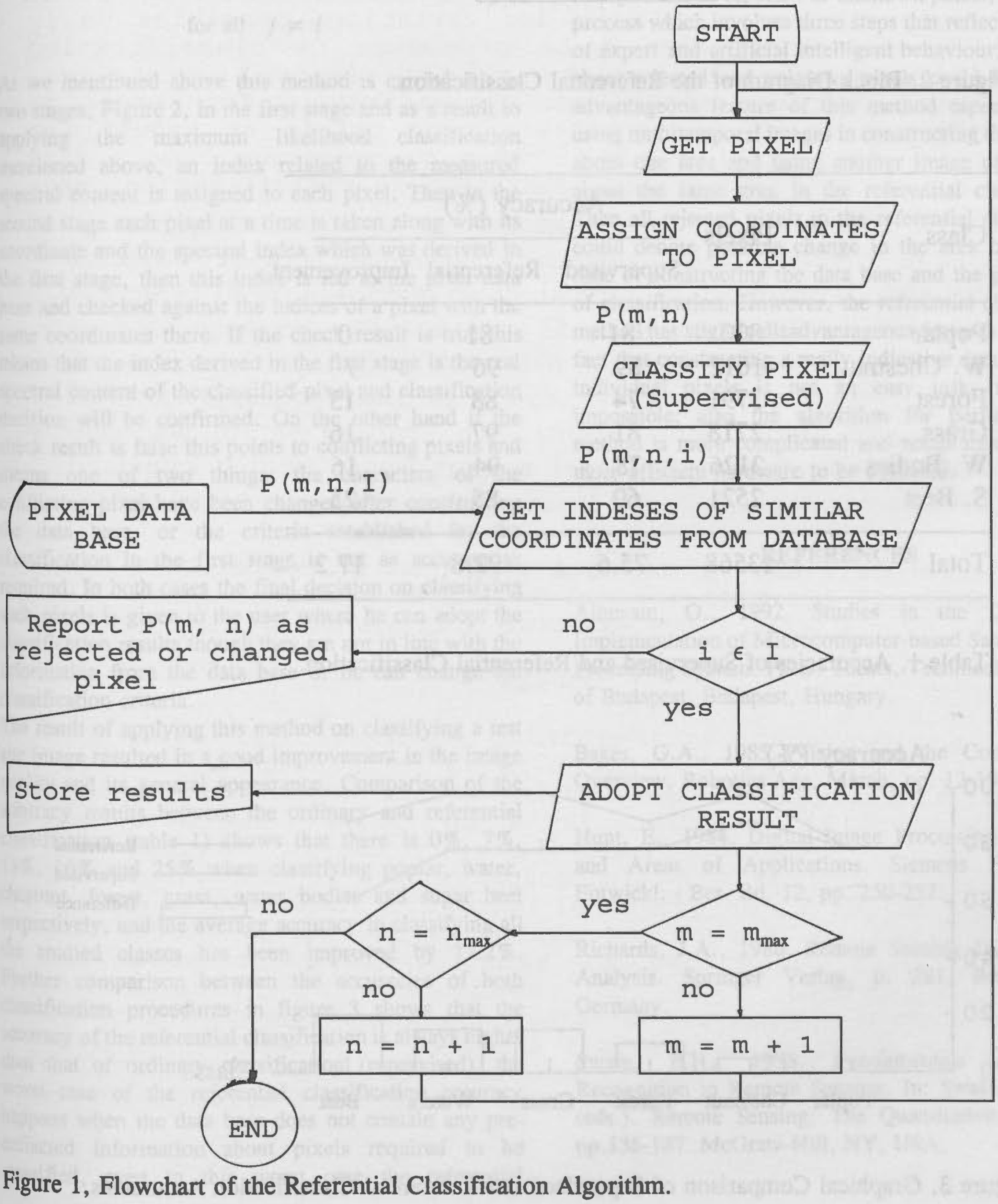


Figure 1, Flowchart of the Referential Classification Algorithm.

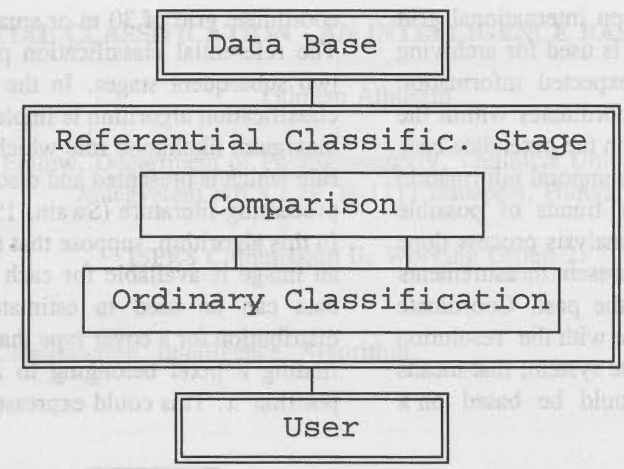


Figure 2, Block Diagram of the Referential Classification.

Class	Pixels	Accuracy (%)		
		Supervised	Referential	Improvement
Poplar	800	81	81	0
W. Chestnut	1612	89	96	7
Forest	3791	74	88	13
Grass	2718	81	97	16
W. Bodies	2126	78	94	16
S. Beet	2521	60	85	25
Total	13568	75.6	92.8	17.2

Table 1, Accuracies of Supervised and Referential Classification.

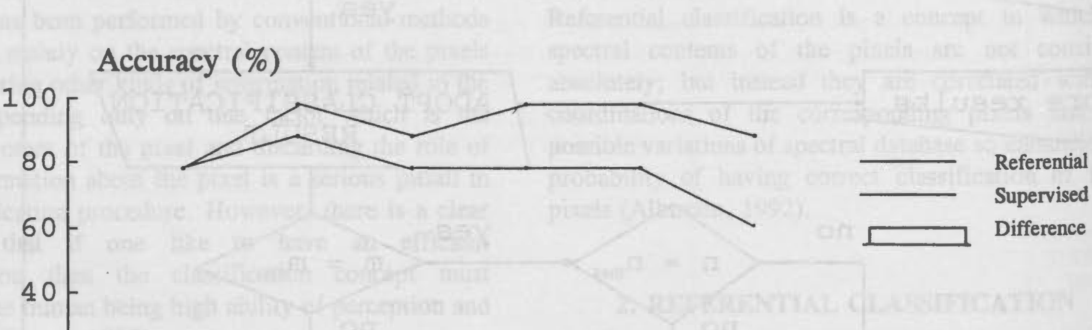


Figure 3, Graphical Comparison of Supervised and Referential Classification Accuracies.

The number of the $p(x/w_i)$ will be equal to the number of the ground cover classes. This means, for a pixel at a position x in multispectral space a set of probabilities can be computed for each class. The required $p(w_i/x)$ of the class and the available $p(x/w_i)$ of training data are related by the Bay's Theorem as follows:

$$p(w_i/x) = p(x/w_i)\, p(w_i)/\, p(x)$$

where $p(w_i)$ is the probability that class w_i occurs in the image. The rule in classifying a pixel at a position x will be:

$$x \ni w_i \text{ if } p(x/w_i)\, p(w_i) > p(x/w_j)\, p(w_j)$$

$$\text{for all } j \neq i$$

As we mentioned above this method is carried out in two stages, Figure 2, in the first stage and as a result to applying the maximum likelihood classification mentioned above, an index related to the measured spectral content is assigned to each pixel. Then in the second stage each pixel at a time is taken along with its coordinate and the spectral index which was derived in the first stage, then this index is led to the pixel data base and checked against the indices of a pixel with the same coordinates there. If the check result is true this means that the index derived in the first stage is the real spectral content of the classified pixel and classification decision will be confirmed. On the other hand if the check result is false this points to conflicting pixels and means one of two things: the characters of the conflicting pixel have been changed after constructing the data base, or the criteria established for the classification in the first stage is not as accurate as required. In both cases the final decision on classifying such pixels is given to the user where he can adopt the classification results though they are not in line with the information from the data base or he can change the classification criteria.

The result of applying this method on classifying a test site image resulted in a good improvement in the image quality and its general appearance. Comparison of the accuracy results between the ordinary and referential classification (table 1) shows that there is 0%, 7%, 13%, 16% and 25% when classifying poplar, water, chestnut, forest, grass, water bodies and sugar beet respectively, and the average accuracy in classifying all the studied classes has been improved by 17.2%. Further comparison between the accuracies of both classification procedures in figure 3 shows that the accuracy of the referential classification is always higher than that of ordinary classification (supervised), the worst case of the referential classification accuracy happens when the data base does not contain any pre-collected information about pixels required to be classified, even in this worst case the referential classification is just as accurate as ordinary classification and is never less as in the poplar class case.

3. CONCLUSIONS

Referential classification of image data is a vital step toward automating the classification process which is an important step in automating the whole image processing and analysis process. This can be achieved by cancelling the role of the user in classifying the conflicting pixels where the spectral data allocated to them in the ordinary classification stage can be adopted, changing their classification to ordinary and not referential or they can be rejected and reported as unknown pixels; a multifold process which involves three steps that reflect high level of expert and artificial intelligent behaviour. Reporting about rejected and unknown pixels could be a highly advantageous feature of this method especially when using multitemporal images in constructing the data base about one area and using another image of later time about the same area in the referential classification. Then all rejected pixels in the referential classification could denote possible change in the area between the time of constructing the data base and the present date of classification. However, the referential classification method has slightly disadvantageous features such as the fact that constructing a really indicative data base about individual pixels is not an easy task, though not impossible; also the algorithm for performing this method is more complicated and needs more time and more efficient hardware to be executed.

REFERENCES

Alhusain, O., 1992. Studies in the Design and Implementation of Microcomputer-based Satellite image Processing System. Ph.D. Thesis, Technical University of Budapest, Budapest, Hungary.

Baxes, G.A., 1985. Vision and the Computer- An Overview. Robotics Age, March, pp. 12-19.

Hunt, E., 1984. Digital Image Processing- Overview and Areas of Applications. Siemens Forsch.- u. Entwickl. - Ber. Bd. 12, pp. 250-257.

Richards, J.A., 1986. Remote Sensing Digital Image Analysis. Springer Verlag, p. 281. Berlin [etc.], Germany.

Swain, P.H., 1978. Fundamentals of Pattern Recognition in Remote Sensing. In: Swain and Davis (eds.). Remote Sensing: The Quantitative Approach, pp.136-187. McGraw-Hill, NY, USA.

CLASSIFICATION RÉFÉRENTIELLE - UN ALGORITHME À BASE D'INTELLIGENCE

RÉSUMÉ

Cet article décrit une méthode récemment mise au point pour la classification des pixels dans les images-satellite. Dans cette nouvelle méthode, appelée «classification référentielle», les contenus spectraux des pixels ne sont pas considérés dans l'absolu; ils sont plutôt mis en corrélation avec un type de données accessoires nécessitant la coordination d'un pixel correspondant et donnant lieu à des variations possibles de la base de données spectrales. Le processus de classification référentielle se déroule ensuite en deux étapes successives. L'application de cette méthode de classification des images de télédétection a donné des résultats plus précis et réduit les interventions de l'utilisateur au cours du processus. On a donc réussi à automatiser davantage l'opération de classification, sans perdre de précision.

AN OVERVIEW OF THE USE OF QUADTREES AND RELATED HIERARCHICAL DATA STRUCTURES IN GEOGRAPHIC INFORMATION SYSTEMS

Hanan Samet

Department of Computer Science
University of Maryland
College Park, MD 20742
U.S.A.

ABSTRACT

Representation is an important issue in geographic information systems. A wide number of representations is currently in issue. Recently, there has been much interest in hierarchical data structure such as quadtrees and octrees. They are compact and, depending on the nature of the region, they save space as well as time and also facilitate operations such as search. In this paper we give a brief overview of the quadtree data structure and related research results. Examples will be drawn from a functioning geographic information system that uses the quadtree as the underlying representation.

UTILISATION DE LA QUADRIPARTITION ET DES STRUCTURES HIÉRARCHIQUES CONNEXES DE DONNÉES DANS LES SYSTÈMES D'INFORMATION GÉOGRAPHIQUE : APERÇU

RÉSUMÉ

Dans les systèmes d'information géographique, la représentation est un point d'intérêt important. On se penche actuellement sur un grand nombre de représentations. Ces derniers temps, on s'est beaucoup intéressé aux structures hiérarchiques des données, comme la quadripartition et l'arbre octogonal. Ces structures sont de faible encombrement et, selon la nature de la région, elles permettent d'économiser temps et espace et facilitent certaines opérations, comme la recherche. Dans cette présentation, nous donnons un bref aperçu de la quadripartition et des résultats de recherches à ce sujet. Nous illustrons notre propos à l'aide d'exemples de systèmes d'information géographique fonctionnels dont le mode de représentation sous-jacent est la quadripartition.

DIGITAL PHOTOGRAMMETRIC SYSTEMS AND THEIR INTEGRATION WITH GIS

IAN DOWMAN

University College London
Gower Street
London WC1E 6BT
United Kingdom

ABSTRACT

Digital photogrammetric systems have been a subject of research and development for a number of years but there is now commercially available equipment and software which has been developed and is supported by major manufacturers. The production of software for extraction of digital elevation models and their use for generating geocoded images is of particular importance. One of the forces behind such development has been the increased need for spatial data for geographical information systems and the use of data from earth observation satellites for this purpose; as a result of this, software packages available for image processing and GIS often include routines for orientation and correction of aerial and satellite images. The result of these developments is increased integration between remote sensing, photogrammetry and GIS.

This paper reviews recent developments in digital photogrammetric systems with emphasis on those which are integrated with GIS. The functionality of the systems will be analysed and evaluated with particular attention paid to the requirements of topographic mapping and GIS. It is important that the user is able to evaluate the quality of the products derived from the digital processing and to monitor each step of the procedure.

Current research will also be reviewed and an assessment made on the progress of development of feature extraction algorithms and the prospects for widespread use of such routines.

KEY WORDS: Digital photogrammetry, automation, Geographical information system, integration.

1. INTRODUCTION

The great increase in the use of Geographical Information Systems (GIS) in recent years has led to a need for consistent, validated data at all scales over wide areas. The combination of digital photogrammetric stereo systems using aerial photographs and satellite data is capable of providing some of that data. The key to producing data of this type, in this quantity, is automation and end to end production systems. At present automation of the mapping process is not automatic and many different systems exist which need to communicate with each other. End to end systems will ensure that data transfer between different stages of production are smooth and that data from various sources can be used. This paper addresses the state of the art in digital photogrammetric systems with emphasis on the problems of automation and integration.

2. REVIEW OF DEVELOPMENTS

The concept of digital photogrammetry was first set out by Sarjakoski in 1981. One of the first systems to be build was the Kern DSP1 which was exhibited at the ISPRS Congress in Kyoto in 1988. It was not until 1992 that systems became available which were suitable for production. One of the first systems was developed by Intergraph as a means of providing data for their own digital mapping system. Other systems arose from defence contracts, particularly for the production of digital elevation models (DEMs). At the same time the use of satellite data has been increasing and image processing systems have been developed to handle this data. in recent years these systems have been coming together with GIS as a step towards integration.

3. CURRENT STATUS OF DIGITAL PHOTOGRAMMETRIC SYSTEMS

3.1 Commercial systems

At the present time there are a number of systems on the market which have been developed to a state useful for production and which are maintained by the vendors. These systems have been described by a number of papers in recent years, for example Dowman et al, (1992) and Heipke, (1993). The distinction between fully functional photogrammetric instruments and software packages designed as tools for GIS or image processing packages for remote sensing applications is still quite marked and the methods for evaluating such systems will differ. Heipke (1993) gives a full list of existing systems with references for further details. The main stereoscopic systems are the Intergraph ImageStation and the Leica DPW 710, 750 and 770 . These offer production of a full range of photogrammetric products: digital mapping, aerial triangulation, DEMs and orthoimages; the image moves relative to a fixed reference mark. The Zeiss Phodis ST has been released in its fixed image moving cursor form and the moving image fixed cursor version will be available later this year.

Less expensive stereo systems such as the Leica DVP and the R-Wel DMS have more limited functionality and accuracy.

A number of systems are available just for the production of DEMs and orthoimages, often these are part of an image processing or GIS. Examples are ERDAS which is an image processing system closely integrated with ARC/INFO, PCI Airphoto ortho and Ramtek HI-VIEW. This type of system is often not fully functional and is designed for the non specialist user. Care should be taken to check the accuracy attained and the algorithms used as these can affect the quality of the final result.

For a fully functional photogrammetric system to be attractive to a production organisation it must offer at least the same functionality and efficiency as an analytical plotter. The user will also expect additional features such as some elements of automation and display. At present the cost and availability of scanning equipment will also be an important factor. For these reasons current equipment concentrates on the production of digital elevation models and orthoimages as these are the areas where automation is most developed and where there is a growing market. Colomer and Colomina (1994) show that there is no economic advantage at the moment to convert to digital systems.

Some of these systems have been developed primarily for use with satellite data and other for aerial photography. Slowly systems are now offering both options.

Under the heading of hardware it must not be forgotten that aerial photographs must be scanned before they can be used in digital systems. Scanners are still expensive and create problems of image quality and data handling. No doubt these problems will be solved in time but at the moment only organisations with large throughputs, such as for orthoimage production, use high precision scanners. A solution for some organisations is the use of scanners designed for the printing trade such as the Sharp JX-600 which has limited resolution (600 dots per inch) instability in the scanning axes, but can nevertheless be used for a number of applications.

Satellite data has been a catalyst to the use of digital data and is one of the main sources of input to digital systems. The principle cartographic product from satellite data is the image map. This has an aesthetic and novelty value as well as being a useful cartographic document and source of data fro map revision. Image maps may be compiled from satellite data which has been warped in two dimensions or corrected by geocoding which involves the use of a digital elevation model (DEM).

The final specification of such products varies according to the producer and the intended market and a wide range exists. Image maps are now derived from commercial production processes. In France Spot Image produces the Geospot range of products through collaboration with IGN Espace, BRGM, ISTAR and Geosys. The IGN Espace concept is described by Galtier and Baudoin (1992). In Sweden, Satelitbild (a SSC company) also has a production flow line and at Institut Cartografic Catalunya (ICC), (Colomina et al 1991) image maps are produced from aerial photographs and satellite data as commercial products.

Line maps from satellite data are less frequently produced and the only well known example is IGN (France). For this standard photogrammetric mapping procedures are used. A number of organisations are developing digital mapping systems which will produce both line maps and image maps, for example at USGS (Skalet et al, 1992) and Canadian Centre for Mapping (CCM) (Ahac et al 1992). A system has been put into production at DMA (Warren 1992). A digital system for use with aerial photographs is in production at the ICC.

3.2 Software

Software for digital photogrammetric systems can be divided into two categories. The first is the orientation and display software, the second is the application software which ranges from programs for collecting digital elevation models through editing collected or imported data to a GIS.

Operations can be automated entirely or in part. An example of partial automation, or computer assisted operations is when the operator sets an approximate position and the final position is fixed automatically. This allows the computer to perform the difficult operation and usually to achieve maximum precision. An example is the line extraction routine in the ISTAR VUE3D system in which the operator roughly indicates the position of a line and the computer finds the exact line of the feature using a dynamic programming technique developed by Mâitre and Wu (1989). The stereo matching of two images to ensure correct height when the operator follows a feature is another example which was introduced in the Kern DSP1 and is now incorporated into a number of systems.

In the production systems automation does not tend to be introduced until a robust algorithm has been developed and proven. Thus most systems still rely on the operator to carry out the inner, relative and absolute orientation with support from image processing routines such as zooming.

Software for the production of DEMs is now is use for production and DEMs can be produced automatically but still need to be edited for blunders and missing areas.

3.3 Research directions

With the exception of the DEM extraction, no commercial systems has a proven automated component. It is evident from the literature that there are a number of areas where automation is seen as having a potential, either in the near or distant future. These areas are:

- identification of ground control points;
- speeding up the orientation process necessary to determine the exterior orientation of the images;
- feature extraction;
- change detection.

3.3.1 Automatic GCP extraction

There is a significant amount of work going on to reduce the dependence on ground control points (GCPs) for absolute orientation and georeferencing procedures.

The identification and selection of GCPs for geometric processing of digital images is usually a time-consuming and expensive process although discussion with producers indicates that this is a small proportion of the total cost of producing image maps. Automatic matching with earlier processed images is a significant help in removing this bottle-neck. This has been discussed and implemented for a number of years, for example Benny (1981). A major problem to be addressed is the accuracy and distribution of extracted points. The problem of map-image matching is much more difficult as work by UCL has shown (Stevens et al 1988).

Building on earlier work (Schenk et. al. 1991) on automatic tie-point determination for the orientation of digital photographs, Toth and Schenk (1992) have described a method of automatically registering images by matching extracted edges and determining identical points.

3.3.2 Automation within the orientation process

In order to determine the orientation elements it is necessary to establish the calibration parameters of the sensor and to fix relative orientation and absolute orientation using ground reference points. The points required for calibration and relative orientation (conjugate points) can already be determined automatically. The fiducial points and the conjugate image coordinates can be derived from stereo matching.

Stokes (1988) developed a fully automated procedure to identify and measure fiducial marks. Fiducial marks are generally different in different cameras. However, they usually have a well defined appearance and occupy an extended area devoid of other information. Their degree of symmetry is high and their approximate location in the image known in advance. They can therefore easily be identified using template matching and then localised with centre of gravity methods.

Haala et al (1993) have described work leading to full automation of the conjugate point problem starting with an interest operator and using an image pyramid to refine the match.

A number of organisations are working with image registration systems which consider whole images (Lee et al, 1993) or layers in images such as roads (ENST in Paris using techniques described by Maitre and Wu, 1989).

Schickler (1992) developed a system for automating the exterior orientation of a single image. It is based on control points which consist of a list of straight 3D-line segments, whose 3D-coordinates are known in a object centred coordinate system, and which mostly represents buildings. Using approximate values for the orientation parameters in order to increase the efficiency, the system matches the images of the object lines with straight line segments, as features extracted from the image. This is done in several steps for all available control points. The final spatial resection is performed with the straight line segments and evaluated with respect to precision and reliability.

This method implies the use of features instead of distinct points for orientation. Orientation in photogrammetry in the traditional approach is more or less completely based on distinct points. However, geometric features in object space can also be straight or curved lines and planar or curved surfaces. These features and their projection into image space can be used for image orientation. Especially for digital techniques of image analysis this aspect is essential, because lines are easier to extract than point features by automatic procedures. In order to rigorously accommodate the extraction and correspondence of these features suitable photogrammetric formulation must be performed. Research and practical work based on the correspondence of general geometric features, beside points, is being performed in computer vision and photogrammetry.

Haala et al (1992) discusses the use of relational matching to match relational descriptions of images and maps. For instance, the top of the seat and the front of the back may be at right angles to each other. The structural descriptions of the images are obtained by thresholding selected channels of colour images and subsequent thinning of the linear structures. The structural descriptions of the landmarks were obtained by digitising maps, but, in principle they could also have been derived from a Geographic Information System (GIS).

3.3.3 Feature extraction

Feature extraction is the subject of intense interest at the moment but except in very well defined areas, there is little prospect of a robust solution to the majority of problems. McKeown, (Dowman et al 1993) notes the major problem to be the complexity and variability of the scene interpretation task. In other words, it is very difficult to design a system which will cope with a wide variety of common scene characteristics. Two major research directions are apparent. The first is the definition of basic concepts and relationships and the design of tools to fit in with these concepts. The second is the use of multiple integration techniques using several different types of data and knowledge based algorithms with a GIS.

The work of Forstner (1992) relies on complex object models, their inter-relationships with their parts and other objects and the variation over time. Molenaar and Fritsch (1991) works along similar lines but has a primary interest in the data within a GIS.

The automated extraction of linear features have been attempted by hierarchical texture analysis (Moller-Jensen 1990), and by search techniques like dynamic programming (Nonin 1992, Gunst and Lemmens 1992, Maitre and Wu 1989) requiring initial approximation of the location of some features (or connections); these initial approximations could come from a GIS or from edges extracted with kernel filters. Nonin (1992) has described a system based on work by Maitre and Wu (1989) and operated by ISTAR, designed for the extraction of linear features. The operator identifies the features approximately and a dynamic programming based algorithm determines their exact position. INRIA are working with IGN (France) and CNES on the problem of extraction of roads. Laser scan are also experimenting with an iterative system based on their VTRAK system.

The classification and segmentation of land use features no longer encourages per-pixel techniques that do not utilise neighbouring information such as texture. There seems to be greater interest in systems utilising *a priori* data from digital maps, GIS, (Janssen and Amsterdam 1991, Bolstead and Lillesand 1992), DEMs and other knowledge sources like human experts (Middelkoop and Janssen 1991). ICC have combined data from a number of sources which include previous classifications, satellite data and topographic maps, the classification is done using neural networks and a 5% improvement on previous results is obtained.

The extraction of buildings require high resolution imagery (e.g.. aerial photographs or possibly the high resolution Russian imagery DD5) and is not achievable by conventional per-pixel classification techniques. Murukami and Welch (1992) have attempted an expert system approach, Gulch (1991, 1992) used a rule-base on initially segmented lines and regions, and supported by consistency checks, but Shufelt and McKeown (1990), for the most successful results so far reported, fused three shape-from-shading and one edge-corner techniques, taking advantage of redundant data and also giving consideration to conflicting information

It can also be concluded that full automation is still a distant goal, computer assistance can play an important role in many applications but the results achieved by the analytical power of the human operator cannot yet be reproduced with sufficient success to remove the operator from the process of feature extraction and object recognition.

4. Integration issues

It is now widely recognised that more information can be obtained from imagery if data from several sources can be used. In order to use such techniques it is essential that these varied images be registered into a single co-ordinate system. Several examples have been quoted above. The image data also has to be used with map data or other data in vector form. To make use of this potentially vast source of data automation has to be used to synthesise the data and extract useful information. For example in the process of extracting features the support of a GIS is becoming essential, a number of methods under development for change detection require the use of existing information to help identify change. In a system for detecting changes in forest under development at UCL under contract to Earth Observation Sciences, edges extracted from imagery are compared to boundaries from maps in order to determine where change has, or has not taken place. In the top down method of feature extraction information about the objects is required and can come from a GIS. Ehlers (1994) identifies a number of issues for investigation:

- data models
- integrated classification procedures;
- integrated error models;
- automated registration and map revision;
- integration with modelling techniques.

These are currently the active areas of research.

5. Validation issues

The concept of accuracy is well established in the evaluation of photogrammetric data. There is no problem in establishing the accuracy of co-ordinates derived from imagery as long as control in the form of check points is available. Moreover the accuracy is usually related to the scale at which a map is produced. Once data from different sources is merged the origin may be lost and if data or images of different scales are merged then assumptions may be made which are not valid. For this reason data should be given an attribute which indicates its source and accuracy. Little effort has been made to do this at present although work has been done to assess problems caused by generalisation at different scales (Jaoa, 1994).

In the area of Radar data which has been geocoded products are produced to indicate certain aspects of accuracy. A pixel area map, also called an energy map, can be produced to indicate the surface area of a pixel when recorded, thus allowing correction for slope effects to be made by the scientist who is interpreting the data. Layover and shadow maps can also be produced (Schrieir, 1993). Error budgets are also produced with geocoded data which include an assessment of the accuracy of the DEM used for the geocoding.

At the Remote Sensing Institute of the European Union's Joint Research Centre at Ispra, a geophysical processor is being developed which aims to provide an application scientist with all the information, and tools to use that information, to allow accurate interpretation of the data.

6. Conclusions

Digital photogrammetric systems are now established as production tools although the economic advantages are not yet well established. Photogrammetric software is also available on digital image processing systems although care must be atken in using this and accuracy assessments made. The use of the latter systems with GIS enables data derived from remotely sensed data to be directly passed to the GIS and this offers significant avantages in data acquisition.

The big break through will be when automation can be applied as a production tool. At the moment this only applies to elevation models and automatic feature extraction and model orientation is still a very active research area.

As these new algorithms are being developed and incorporated into digital systems it is important that quality is considered and validation techniques built in.

References

Ahac A. A., Defoe R., Wijk M. C. van (1992); 'Considerations in the Design of a System for the Rapid Acquisition of Geographic Information', PE&RS, Vol. 58, No. 1.

Benny A. H. (1981); 'Automatic Relocation of Ground Control Points in Landsat Imagery', Paper presented at a RSS society conference on "Matching remote sensing technologies and their applications", London, 16-18 December, 1981.

Bolstead P. V., Lillesand T. M. (1992); 'Rule-Based Classification Models; Flexible Integration of Satellite Imagery and Thematic Spatial Data', PE&RS, Vol. 58, No. 7, pp. 965-971.

Colomer J and Colomina I, 1994. Digital Photogrammetry at the ICC. paper accepted for publication in Photogrammetric Record.

Colomina J, Navarro J, Torre M, 1991. 'Digital Photogrammetric Systems', Digital Photogrammetric Systems at ICC. Wichmann Verlag, Karlsruhe. Pages 217-228.

Dowman I.J., Ebner H, Heipke C, 1992. Overview of European developments in digital photogrammetric workststations. PERS 58(1):51-56.

Dowman I. J., Muller J-P., Nwosu A. G. (1993). A report on developments in the field of automated catography. Contract report, DRA Contract CBFRN1B/182. 127 pages

Forstner W. (1992); 'Feature Extraction for Digital Photogrammetry. Photogrammetric Record 14(82)595-611.

Galtier B and Baudoin A, 1992. Image maps.Proceedings pf ISU Working Group Session. ISPRS Congress, Washington DC, August 1992. Pages 74-77.

Gulch E. (1992); 'A Knowledge Based Approach to Reconstruct Buildings in Digital Aerial Imagery', Proceedings, ISPRS Congress, Washington D.C., August 1992, Comm. II, pp. 410-417.

Gulch E. (1991); 'Using a Knowledge Based System in the Generation of a Generic Model for Buildings', Proceedings IGARSS '91, Espoo, Finland.

Gunst M. E. de, Lemmens M. J. P. M. (1992); 'Automated Updating of Road Data Bases From Scanned Aerial Photographs', International Archives of Photogrammetry and Remote Sensing, 29(IV) :477-484.

Haala B, Hahn M, Schmidt D, 1993. Quality And Performance Analysis Of Automatic Relative Orientation. Integrating Photogrammetric Techniques With Scene Analysis And Machine Vision. SPIE proceedings 1944. pp 141-150.

Haala, N., Vosselman, G., 1992. Recognition of road and river patterns by relational matching. IntArchPhRS 29(B3), pp 969-975.

Heipke C, 1993. Performance and state-of -the-art of digital stereo processing. photogrammetric Week '93. Wichmann. Pp173-183.

Janssen L.L.F, Amsterdam J. Van (1991); 'An Object-Based Approach to the Classification of Remotely Sensed Images, Proceedings, IGARSS-Symposium 1991, Espoo, Finland.

Jaoa E, 1994. Causes and consquences of map generalisation. PhD Thesis, University of London.

Lee A J, Carender N H, Knowlton D J, 1993. Fast autonomous registration of Landsat, SPOT and ADRG map data. Proceedings of SPIE Conference 1944, Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision, Orlando, April 1993.

Maitre H., Wu Y. (1989); 'IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 37, N0. 2.

Middelkoop H., Janssen L. L. F. (1991); 'Implementation of Temporal Relationships in Knowledge Based Classification of Satellite Images, PE&RS, Vol. 57 No. 7, pp. 937-945.

Molenaar M., Fristch D. (1991); 'Combined Data Structures for Vector and Raster Representation in GIS'; Vol. 4, No. 5, Karlsruhe, 1991, pp. 26-32.

Moller-Jenssen L. (1990); 'Knowledge Based Classification of an Urban Area Using Texture and Context Information in LANDSAT-TM Imagery, PE&RS, Vol. 56 No. 6, pp. 899-904.

Murakami H., Welch R. (1992); 'Automatic Feature Extraction for Map Revision', International Archives of Photogrammetry and Remote Sensing, 29(IV) :569-575.

Nonin P. (1992); 'Computer Assisted Production of Planimetric Data: An Industrial Approach', International Archives of Photogrammetry and Remote Sensing, 29(IV):342-345.

Sarjakoski T. (1981); 'Concept of a Completely Digital Stereoplotter', The Photogrammetric Journal of Finland, 8 (2), 95-100.

Schenk T., Li J. C., Toth C. K. (1991); 'Towards an autonomous system for the orientation of digital stereo-pairs', PE&RS, 58(8):1057-1064.

Schickler, W., 1992. Feature matching for outer orientation of single images using 3-D wireframe controlpoints. IntArchPhRS, 29(B3): 591-598.

Schreier G, 1993. SAR Geocoding:Data and Systems. Wichmann.

Shufelt J. A., McKeown D. M. (1990); 'Fusion of Monocular Cues to Detect Man-Made Structures in Aerial Imagery', Report manuscript of the Computer Science Department, Carnegie Mallon University, Pittsburg, USA; a revised and extended version of a paper presented at the IAPR Workshop on Multisource Data Integration in Remote Sensing, College Park, MD, USA, June 14-15, 1990; 31 pages.

Skalet C. D., Lee G. Y. G., Ladner L. J. (1992); 'Implementation of Softcopy Photogrammetric Workstations at the US. Geological Survey', PE&RS Vol. 58, No. 1, pp. 57-63.

Stevens A., Morris A. C., Ibbs T. J., Jackson M. J., Muller J-P. (1988); 'Automatic Generalisation of Image Ground Control Features From a Digital Map Database', International Archives of Photogrammetry and Remote Sensing, 27(II/3):402-413.

Stokes, J., 1988. A photogrammetric monoplotter for digital map revision using the image processing sytem GOP-300. IntArchPhRS, 27 (B2):424-433.

Toth C. K., Schenk T. (1992); 'Feature Based Matching for Automatic Image Registration', ITC Journal, 1992-1, pp. 40-46.

Warren L. R. (1992); 'Digital Production System', Proceedings ISPRS Congress, Washington D.C. 1992, Comm. II, pp. 631-635.

Les systèmes de photogrammétrie numérique et leur intégration au SIG

Résumé

Pendant nombre d'années, les systèmes de photogrammétrie numérique ont fait l objet de recherches et ont connu certains perfectionnements. On trouve maintenant sur le marché du matériel et des logiciels conçus par les principaux fabricants, qui en assurent également le soutien technique. La production de logiciels d'extraction de modèles numériques d'altitude, qui permettent de produire des images géocodées, revêt une importance particulière. Les motifs qui sous-tendent ce progrès sont, entre autres, le besoin croissant de données spatiales, qui servent à alimenter les systèmes d'information géographique, et l'utilisation à cette fin des données des satellites d'observation terrestre. Pour cette raison, les progiciels de traitement d'images sur le marché et les SIG comprennent souvent des sous-programmes d'orientation et de correction des photographies aériennes et des images obtenues par satellite. Ces perfectionnements ont permis d'améliorer l'intégration de la télédétection, de la photogrammétrie et des SIG.

Cette présentation porte sur les progrès récents des systèmes de photogrammétrie numérique et met l'accent sur ceux qui sont intégrés à des SIG. On y analyse et évalue la fonctionnalité des systèmes tout en accordant une attention particulière aux exigences en matière de SIG et de cartographie topographique. Il est important que l'utilisateur soit en mesure d'évaluer la qualité des produits obtenus par traitement numérique et qu'il puisse contrôler chaque étape du processus.

On examine également l'état actuel de la recherche et on évalue les perfectionnements au titre de la conception d'algorithmes d'extraction des caractéristiques, de même que les possibilités d'en répandre l'utilisation.

THE USE OF SPATIAL INFORMATION OF CANADA POST CORPORATION

Ken Tucker

abstract

Not available at time of printing

Utilisation des données à référence spatiale de la Société canadienne des postes

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif

MANAGEMENT PERSPECTIVE OF AN INFRASTRUCTURE FOR GIS INTEROPERABILITY - THE DELTA-X PROJECT

MOSAAD ALLAM

Geographic Information Systems Division
Surveys, Mapping and Remote Sensing Sector
Natural Resources Canada
615 Booth Street, Ottawa, Ontario, Canada K1A 0E9

ABSTRACT

Shared access to a federated set of dispersed databases located at the nodes of a wide area network is a major issue. Considering that these nodes themselves are part of a local area network and the possibility that there can be a variety of GISs and DBMSs residing on a heterogeneous array of platforms, this precipitates complex problems of data transfer from system to system, not to mention problems arising from the dissimilarities in format and the standards for interchange. From the users point of view, the data should be current, available in readily usable format, easily accessible, supported with database management functionalities and allow selective feature retrievals. The user should have the option to use whatever database or GIS which is appropriate for the application.

Delta-X is an implementation of a federated multi-database environment consisting of a wide area network of servers connected by high-speed communication lines. This paper addresses the management in developing such a system. Interesting problems of design, prototyping, testing and design revisions are addressed. In particular the lessons learned in the design, revision, tracking and technology that impact the release of the system, as well as shifts in software development paradigms that enhance the development of the system.

INFRASTRUCTURE PERMETTANT L'INTEROPÉRABILITÉ DES SIG - PROJET DELTA-X : LE POINT DE VUE DE LA GESTION

RÉSUMÉ

L'accès partagé à un ensemble fédératif de bases de données dispersées aux noeuds d'un réseau à grande distance constitue un problème majeur. Si l'on considère que ces noeuds font eux-mêmes partie de réseaux locaux ainsi que la possibilité que toute une gamme de SIG et de SGBD puissent résider sur un ensemble hétérogène de plates-formes l'on peut imaginer toute la complexité du transfert de données d'un système à un autre sans parler des problèmes suscités par des différences entre les formats et les normes en matière d'échange de données. Du point de vue des utilisateurs, les données devraient être à jour, présentées suivant un format utilisable, d'accès facile, être explorables par l'entremise d'un système de gestion de base de données et permettre l'extraction sélective d'éléments. L'utilisateur devrait avoir le choix d'utiliser la base de données ou le SIG qui conviennent pour l'application à laquelle il les destine. Le Delta-X est une mise en oeuvre d'un système fédératif de bases de données multiples intégrant des SIG et des bases de données dans un environnement uniforme consistant en un réseau à grande distance dont les serveurs sont reliés par des lignes de télécommunication à grande vitesse. Dans la présente communication l'on traite des questions de gestion dans la mise au point d'un tel système. D'intéressants problèmes de conception, de mise au point de prototypes, de mise à l'épreuve et de révision des plans sont abordés. Les leçons tirées en matière de conception, de révision, de suivi des changements technologiques influençant la diffusion du système et de modification des paradigmes de mise au point de logiciels améliorant le développement du système sont présentées.

1. INTRODUCTION

In 1987, when the GIS Division, Natural Resources Canada were set-up, our strategic objectives was to develop applications to promote the use of Sector data, demonstrate the use of GIS technology, improve access to data through improved communications facilities, and to explore research that will aid in the growth of GIS technology. Early 1988, a project was initiated to develop the National GIS Technology Centre in the Division. Two basic guide lines were established. First, the GIS Technology Centre would include a variety of GIS technologies to avoid promoting one vendor or technology over the other and to give our staff the opportunity to build expertise in many areas. All GIS workstations, computer, and peripheral devices would be set-up in a local area network (LAN).

The division acquired several GIS software packages, namely: SPANS from TYDAC, CARIS from USL, ARC/INFO from ESRI Canada, Vision from GeoVision (currently SHL, Ottawa), Microstation from Intergraph Canada, Horizon from Laser Scan, PAMAP from Essential Planning Systems Ltd., GRASS by the U.S. Army Corps of Engineers, IDRISI from Clark University, Easi-Pace from PCI, TIMS from MLA, and MAP-INFO. Other data input software e.g., the raster/vector conversion software CAD-CORE by Hitachi, and V-TRACK by Laser Scan were also acquired. This software was installed on platforms, ranging from Personal Computer (PC) based systems, through systems that could work on any general Unix workstation environment to specialized GIS workstations such as the Intergraph MicroStation. The various systems were interconnected into a Unix local area ethernet network. The operation of this simple LAN based system quickly revealed a number of operational problems.

The project in 1988, by consulting with various users within our Sector, and other department at the federal, provincial levels of governments, suggested the idea of GIS interoperability and the need for communications between multiple GISs in a heterogeneous environment. In addition, we initiated several studies on hardware requirements, the selection of a suitable DBMS for the project implementation, and in 1989 completed the design of the first system prototype. Over a period of one year the first prototype was developed. Early in 1991, our project team started the design and implementation of the current version of the system.

This paper describes the operational framework for the development of Delta-X and a Spatial Browser, the management perspectives and technology support, project management, monitoring and control. Issues related to resources, testing and evaluation are briefly discussed.

2. PROJECT OBJECTIVES

The primary objectives of the projects are:

- to develop a federated multi-database management system that provides interoperability between different GISs : DBMS' and/or formats (the Delta-X); and;
- to develop a spatial data browser as a front end to the Delta-X, that facilitates the access to metadata of various databases, e.g.: information on specific datasets, ownership, geographic coverage, format, availability, access mode, cost, etc.

2.1 System Functionalities

The Delta-X system functionalities are separated into two: the server functionalities and the client functionalities. The clients serves both as a database client to the database at the default attached server and as a server for some specific vendor GIS. The key requirement of the Client Interface is that it must be configured as an X11/R5 server. The client is a X11 server running either Motif or an OpenWindow interface and is capable of displaying the graphics from a remote X client (i.e., a DxServer).

From a Client workstations, at which a user runs some GIS software, the user is able to launch the Spatial Browser to identify a dataset of interest. The search criteria for a dataset is by keywords, project titles, data-set owners, area or coverage, etc. An interesting feature of the Spatial Browser is that it can allow spatial searches by invoking a map and selecting the region on the map to constrain the search. Facilities exist from the Spatial Browser to invoke Delta-X, and make off-line (E-mail, fax) requests for some identified information relevant to the user's project.

Invoking Delta-X from a user workstations first establishes connection to a default DxServer. Delta-X provides functions for:

- Displaying selected or all features of a vector map, the image or even a digital elevation model of a coverage.
- Retrieving a specific data-set, either vector, raster, etc. and translating to the user's specific format.
- Administrative functions such as registering and de-registering a database, starting up the dxServer after failure, authorizing users, etc.
- Monitoring the progress of data-being retrieved and the state of a ongoing transaction.

Although Delta-X is not a Geographic Information System, it has the potential of being used in conjunction with other user libraries and software for full GIS application development. In particular, it provides the opportunity for those interested in exploiting parallel and distributed processing capability to develop algorithms for their specific applications.

3. OPERATIONAL FRAMEWORK

The work on this project stared with an evaluation of the computing systems, GIS software packages and database management technologies. There were a number of emerging technologies and we opted for Commercial of the Shelf Components. The functionalities of several DBMS' were studied and Ingres RDBMS was selected. A spatial topological data model was developed and implemented.

The environment we considered for the system implementation included:

- large organizations which often use different (and incompatible) GISs to store their spatial data in different databases, further GISs are nodes connected to a Delta-X server in a LAN,
- a network of wide area networks (WAN) connecting multiple Delta-X servers; and
- individual users (Delta-X clients), who are authorized to connect over a LAN or a WAN to at least one Delta-X server.

A Heterogeneous GIS environment was selected due to the wide variety of GIS software and hardware platforms available on the market. The approach we took in our solution was to consider the databases residing in various agencies to be autonomous. The availability of autonomous databases over a WAN to users has many implications. First, the dataholders must be part of a federated environment and the data users must have authorization to access data in the federation. The federation must allow browsing of databases so that users can identify the data required and request transfers. Second, the data holders manage and control their spatial data and metadata, so that data will be up to date, consistent, error free and secure from unauthorized changes. In this environment, the user can browse the metadata database for data sources. Data transfer from "donor' to user 'target' systems is executed by the Delta-X.

4. MANAGEMENT PERSPECTIVE AND TECHNOLOGY SUPPORT

One of the cost effective methods the Division used to see if the users' needs are satisfied was to implement a research prototype system based on the requirement analysis and to test the functional capabilities of the design. This way, the concepts that were established earlier could be tested on a real system to make sure they work. This allowed our team to install the prototype on a local area network (LAN), make effective assessment of the development, conduct performance analysis and establish requirements such as storage and display capability. The results of these tests made it possible for further system enhancements towards developing an industrial strength prototype system that will meet the challenges of data exchanges over a wide area network (WAN) without rehashing much of the code developed for the research prototype.

Moreover, the rationale that was followed was to use existing technology as much as possible. To illustrate this in the context of Delta-X, we decided to use PHIGS, RPC, Open Windows, MOTIF and INGRES among other products. The integration of these technologies into Delta-X was necessary for the effective performance of the system. The conformance to such standard packages helped the system designers to make it highly portable.

5. RESOURCES

A project of this nature needs a funding strategy to encourage the developers to reach the goals and even surpass them. The limited funding forced the Division to choose only the essential elements to be implemented in the initial phase leaving further development subject to the availability of funding. As you will find out from the papers in this session (Otoo, 1994), (Medved, 1994) and (Chaly, 1994), astounding developments have taken place and a unique solution has been found for the exchange of digital data in a federated multidatabase environment.

6. PROJECT MANAGEMENT, MONITORING AND CONTROL

The project was configured as a full target system including all the essential components. The system was designed as top down system to provide a sensible working system. All the necessary connectivity was established. At this point a bottom up approach was adopted for the implementation and integration of the system components. The requirements for integrating GIS vendor products and establishing translators became quite apparent.

Normally, before starting a project of this magnitude, it is essential to establish test suites to ascertain whether the system can perform according to the design criteria. We relied on simple test criteria in our limited funded initiative. For example, the data available in CCOGIF format was displayed in one of the GIS systems and compared to the same data translated to Delta-X and transferred to the same GIS. In addition the data within Delta-X was displayed using PEX. This three way test provided confidence in the reliability of the data translation and conversion process. Developing comprehensive test procedures is in progress to assess the accuracy and reliability of network transaction and other features of the system. Data transfer between Delta-X clients with different GISs, was undertaken via our SUN 690 server, and communication between our Delta-X server and another server was also successfully tested.

7. TESTING AND EVALUATION

The success of the implementation of Delta-X and GIS Spatial Browser can be attributed to the stability and the reliability attached to the Alpha release which is one step ahead of the prototype. The successful integration with GIS software and the adaptability to technology changes can be seen in the alpha release, if you consider that the project started in 1989 when this latest technology was not available. The system is capable of integrating new versions of existing GISs and adopting new protocols and standards. Today, we are witnessing high speed network interconnectivity and the shrinking of the gap between LANs and WANs. These are technological advances that will certainly have an impact on the final system implementation.

The evaluations performed on the system have enabled us to determine whether the system meets the performance criteria and to develop techniques to tune the system to achieve the design criteria. Both theoretical and practical analysis were conducted to re-evaluate user requirements and introduce new ideas and features through revision and reengineering.

8. MANAGEMENT FRAMEWORK: BEYOND TECHNOLOGY

The administration of a system such as Delta-X requires a management framework. For example, if data federation is implemented, a number of policies have to be established to control the federated environment. Policy such as authorization, security, charges for data transfer and many others require management decisions. For example, it may be that the data holder prefers to put a disclaimer to avoid any legal implications in the public use of the data. These are issues and policies for senior management to decide on.

9. FUTURE IMPLEMENTATION STATUS

The implementation of Delta-X is still in progress. However a substantial portion of the system is operational. The current implementation runs three servers and one Delta-X master. These use INGRES and ORACLE database management as the underlying databases.

All server currently run some flavour of the UNIX operating system and are interconnected using ethernet. All remote communication, either between servers or between a server and client uses SUN remote procedure calling over TCP/IP protocols The common query language for remote data access is SQL. This is restructured to the SQL flavour of the host DBMS of the DxServer.

While the development has been challenging both from research and academic point of view, its main motivation has been the desire to solve a common problem identified in a number of organizations and institutions having different GIS technologies. This is the problem of an integrated access of databases and data-sets in a heterogeneous computing network.

10. CONCLUSIONS

Although the Delta-X prototype works with GIS data, its architecture, design and implementation can be applied to any type of data. The successful implementation of the Delta-X and Spatial Browser in our Division, and the unprecedented capabilities offered to all the GIS workstations on our network. They can browse through the directory and metadata database of the federal government data holdings developed by the Inter Agency Committee on Geomatics. Data transfer between the various GIS databases and standard file formats are further proof of the validity of our concept and the success of our implementation. The unveiling of the prototype is scheduled for the ISPRS Commission II Symposium and GIS'94 Conference, June 1994.

REFERENCES

Otoo, E., and Allam, M., 1991, A Spatial Topographical Data Model: A FirstStep Towards Standardisation of GIS Data Exchange. Proceedings, The Canadian Conference on GIS 91, Ottawa, Ont., Canada, March 18-22, 1991, p. 763.

Medved, J., Petras, J., 1994, The Client-Server Architecture of Delta-X, Proceeding, ISPRS Commission II Symposium, Ottawa, Ont., Canada, June 6-10, 1994, pp.

Otoo, E., 1994: A Federated Spatial Information System: Delta-X, Proceeding, ISPRS Commission II Symposium, Ottawa, Ont., Canada, June 6-10, 1994, pp.

Chaly, C.K., Zhu, W., and Effah, S., 1994, MetaView: A GIS Spatial Browser - Functions and Services, Proceeding, ISPRS Commission II Symposium, Ottawa, Ont., Canada, June 6-10, 1994, pp.

Infrastructure permettant l'interopérabilité des SIG - Projet Delta-X : Le point de vue de la gestion

RÉSUMÉ

L'accès partagé à un ensemble fédératif de bases de données dispersées aux noeuds d'un réseau à grande distance constitue un problème majeur. Si l'on considère que ces noeuds font eux-mêmes partie de réseaux locaux ainsi que la possibilité que toute une gamme de SIG et de SGBD puissent résider sur un ensemble hétérogène de plates-formes l'on peut imaginer toute la complexité du transfert de données d'un système à un autre sans parler des problèmes suscités par des différences entre les formats et les normes en matière d'échange de données. Du point de vue des utilisateurs, les données devraient être à jour présentées suivant un format utilisable, d'accès facile, être explorables par l'entremise d'un système de gestion de base de données et permettre l'extraction sélective d'éléments. L'utilisateur devrait avoir le choix d'utiliser la base de données ou le SIG qui conviennent pour l'application à laquelle il les destine. Le Delta-X est une mise en oeuvre d'un système fédératif de bases de données multiples intégrant des SIG et des bases de données dans un environnement uniforme consistant en un réseau à grande distance dont les serveurs sont reliés par des lignes de télécommunication à grande vitesse. Dans la présente communication l'on traite des questions de gestion dans la mise au point d'un tel système. D'intéressants problèmes de conception, de mise au point de prototypes, de mise à l'épreuve et de révision des plans sont abordés. Les leçons tirées en matière de conception, de révision, de suivi des changements technologiques influençant la diffusion du système et de modification des paradigmes de mise au point de logiciels améliorant le développement du système sont présentées.

Mots clés : Fédération, SIG, SGBD, base de données.

The Client-Server Architecture of Delta-X

Jan Medved
Natural Resources Canada, 615 Booth Street, Ottawa, Ontario, K1A 0A9

Juraj Petras
Interobjects Spatial Research, P. O. Box 5836, Ottawa, Ontario, K2E 6K2

Abstract

Delta-X is a federated multidatabase system that provides inter-operability between different database management systems (DBMS). It is based on the assumption that each underlying DBMS is based on the client-server architecture and each client workstation is connected to a network that is configured with access to at least one DBMS server. A client, besides being able to query and access a server database, is typically configured as a geographic Information System's workstation. This paper describes the separate functionalities of a client and a server, and the mode of interconnection and inter-operability between them. We describe the architecture, configuration requirements, services offered by servers and client-server and server-server communication within the Delta-X system.

Keywords

Distributed databases, inter-operability, transaction processing systems, client-server architectures, GIS.

1.0 Introduction

The deployment of large local and wide area networks (LANs and WANs) that connect thousands of client and server nodes together with the availability of a great variety of software packages that store and manipulate data (in our case, geomatic data) will soon cause software interoperability and data compatibility problems. Different people and departments within large organizations, which often use different (and incompatible) GIS systems and store their geomatic data in different locations, need to share their data with each other. Also, organizations need to share data with other organizations.

Apart from the software and data compatibility problems, sharing of geographically distributed data poses multiple organizational problems, such as how to advertise the availability of data to the user community, or how to store data for best availability and/or transfer performance.

To address the above issues, we designed and implemented Delta-X, a system that connects over a LAN or WAN a heterogeneous network of databases servers to various GIS clients. A GIS client can be a workstation running any commercially available GIS system, such as CARIS [6] or ARC/INFO [5]. Delta-X creates a platform for a seamless exchange of geomatic data between different systems on different locations. Delta-X enables uniform integrated access to data from different geographic information systems. A GIS can export data to Delta-X or import data from Delta-X in many different formats, such as CCOGIF, SIF, ARC/INFO, or CARIS. Our current implementation of Delta-X stores different types of geomatic data - spatial data (vector or raster), spatial data attributes, structured text, and others. This can be easily modified to store other types of spatial and non-spatial data.

From a user's perspective, Delta-X is a distributed billboard through which users can exchange data. Data enters Delta-X in one format on one location, and the same data exits Delta-X in a different format on a different location. Note also, that although the Delta-X prototype works with GIS data, its architecture, design, and implementation can be applied to any type of data. In a sense, GIS data constitutes a very difficult case among applications that can use the Delta-X paradigm because of large volumes of data that need to be stored, converted, and transferred within a network, and because of a large variety of formats that the data can be presented in.

The underlying paradigm of Delta-X is that of a distributed database. Data in Delta-X is stored on multiple servers distributed across a LAN or WAN. The paradigm of Delta-X is also that of a federated database. That means, servers in a federation form a community which another servers can join, or which the servers in the community can leave. When a server joins a federation, it makes all or some of its data available to other servers and clients in the federation. In turn, it also gains the right to access data on other servers. Access to data on servers within the federation by clients is subject to authentication procedures.

This paper examines architectural and implementation issues related to Delta-X, focusing on the design of the Delta-X server and on the design of the Delta-X client for remote operations and for controlling remote transactions. The paper is organized as follows. Section 2.0 describes the basic architecture of Delta-X and gives the rationale for its design decisions. Section 3 describes the Delta-X server and its services. Section 4 describes the Delta-X clients and their services. Section 5 compares Delta-X to other client/server distributed data management and movement architectures. Future work is described in Section 6. We conclude with Section 7.

2.0 The Architecture of Delta-X

2.1 Requirements and Functional Specification

Delta-X is intended for two types of applications: First, it can be used in large geographically distributed organizations in which different departments share data with each other and which use a multitude of mutually incompatible software packages and/or machines. An example of such organization is a government with its multiple levels (federal, provincial, and municipal).

Second, Delta-X can be used in organizations which sell data and their customers.(e.g. DigiMap) A data producer organization can consolidate their data management and storage and reach the widest possible customer base (which uses a variety of different GIS systems running on different machines). Such organization can distribute the Delta-X software among its customers, and store the software on one or more Delta-X servers. A data consumer organization on the other hand, can obtain data from any data producer in a format that can be used by the software deployed in the organization.

In both types of applications, each shared data has one owner and data is typically updated only by the owner (which is the producer of the data), but concurrently read-accessed by many consumers of the data. Most transactions in the system are read accesses to the shared data which do not update the data or the state of the system. For write accesses to data, the system must support data check-in and check-out - data is checked out for an update (which can take several weeks to complete) and checked back in after the update has been finished.

The system must support a high degree of concurrency, as many transactions accessing the same data can be active concurrently. The system must also give the end-user the highest possible degree of control over a transaction - a transaction must "feel" as if it would execute on the end-user's machine. On the other hand, the requirements for roll-back and recovery are not so stringent - if a transaction fails, the end-user will restart it, and the system must only recover used-up disk space. Recovery means to restart a transaction (from a given check-point).

To satisfy the above requirements, Delta-X must handle three main tasks. First it must handle the storage of the shared data. Second, it must handle the conversion of shared data, stored in an internal data interchange format, into all desired GIS formats so that the data can be used by different target systems. Delta-X also must handle the conversion of data from target systems into the internal data interchange format. Third, Delta-X must facilitate the transfer of converted data from Delta-X to a target system and the transfer of data to be converted and stored in Delta-X from a target system to Delta-X.

A universal data interchange format makes Delta-X easily extensible: each new GIS data format only requires two new conversion routines - to and from the data interchange format. Without the data interchange format, adding a new data format from a new GIS to Delta-X would require to implement conversion routines to and from all formats already supported by Delta-X (Supporting n different data formats would require n^2 conversion routines).

2.2 The Architecture

The architecture that handles all the above tasks is shown in Figure 1. Delta-X handles the above tasks, performing multiple client-server roles [4]. First, Delta-X servers control the data storage in commer-

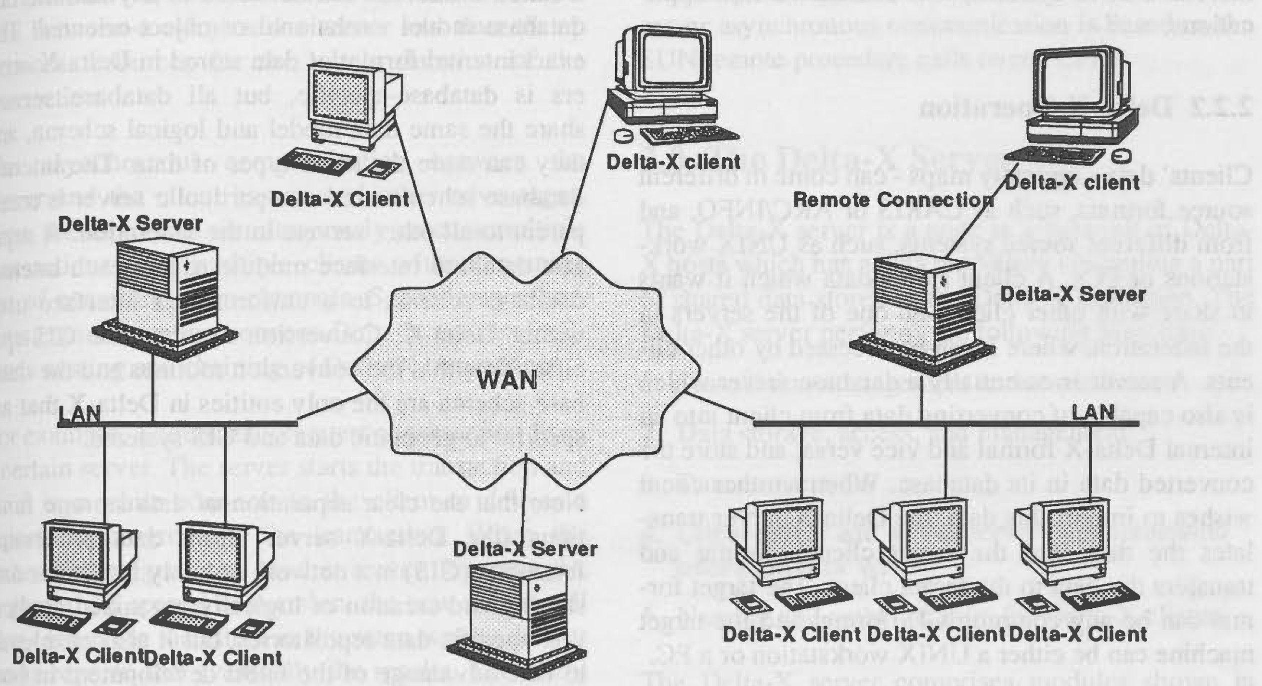


Figure 1: The Architecture of the Delta-X System

cial DBMS systems. Second, Delta-X is a client-server transaction processing system - it handles message and data exchange between Delta-X database and/or other specialized servers and Delta-X clients. Third, Delta-X performs specialized servers functions, such as data conversions to and from the internal data interchange format. Finally, Delta-X acts as a proxy client on behalf of GIS systems that produce or consume data - it is the source or destination of data in Delta-X transactions.

2.2.1 Delta-X Structure

As can be seen in Figure 1, Delta-X is a loosely coupled network of Delta-X servers and Delta-X clients. Clients and servers attached to the same LAN form a *cluster.* Clusters are connected to each other via a WAN, which forms the backbone of the Delta-X system. Servers and/or clients can also be connected directly to the backbone WAN or, via a dial-up line, to one of the cluster servers. The set of all clusters, clients and servers connected directly to the backbone WAN, and clients remotely connected to cluster servers forms the *Delta-X federation.* Communication between clients and servers in the federation is IP-based (both over the LAN and the backbone WAN). Any client can request a transaction from any server in the federation.

A dedicated server in a cluster performs name and authentication services for all clients in the cluster. A server can also perform name and authentication services for clients remotely attached to the server. Any server on the cluster can cache data from remote servers. Although having a dedicated server for certain functions, a client can connect to and request data from any server in the federation.

A complementary system, the MetaView spatial browser [2] stores the meta-data on the databases in the federation. However, MetaView's function is not limited to Delta-X. A special server, the DxMaster monitors and coordinates the activities of servers in the fedration. When a server wants to join or leave the federation, it contacts the DxMaster, which updates its configuration and broadcasts the configuration tables to the rest of the servers in the federation. The DxMaster is duplicated for high availability.

Delta-X takes the three-ball approach to transaction processing [3], where presentation services, transaction workflow, and applications are run in three separate processes, usually depicted as balls. The presentation services are performed by Delta-X clients. Transaction workflow is controlled by dedicated modules in Delta-X servers. Applications are the Delta-X data repositories (built on top of commercial DBMS systems) and data conversion applications.

2.2.2 Delta-X Operation

Clients' data - typically maps - can come in different source formats, such as CARIS or ARC/INFO, and from different source systems, such as UNIX workstations or PCs. A client stores data which it wants to share with other clients on one of the servers in the federation, where it can be accessed by other clients. A server is essentially a database server which is also capable of converting data from client into an internal Delta-X format and vice versa, and store the converted data in its database. When another client wishes to import this data, the Delta-X server translates the data into the target client's format and transfers the data to the target client. The target format can be any common GIS format and the target machine can be either a UNIX workstation or a PC.

A server handles requests both from clients and from other servers in the federation. A transaction starts on one server. However, it may migrate to another server to complete a conversion from one data format into another, or to finish a data transfer.

When a GIS wants to import or export data to Delta-X, it relinquishes control to the Delta-X client. The Delta-X client initiates the data transfer/conversion transaction. During data import operations, the GIS will use the data imported in the transaction after a transaction has been completed.

2.2.3 Delta-X Clients and Servers

Delta-X has a client component and a server component. On a client, Delta-X provides a GUI that enables GIS users to start and control transfer/conversion transactions and communicates with the server. On a server, Delta-X provides the client-server and server-server communication software, a module that transfers data, a network management agent, an interface to a database, and a set of data conversion routines. The server's network management module facilitates setting of server's operational parameters, authentication keys, the registration and de-registration of a server in the Delta-X federation. It also raises alarms to network management when the server malfunctions. A server can leave the federation only if there are no transactions in progress on the server.

Data on the server can be stored in any commercial database model - relational or object-oriented. The exact internal format of data stored in Delta-X servers is database-specific, but all database servers share the same data model and logical schema, and they can store the same types of data. The internal database schema used at a particular server is transparent to all other servers in the federation. A separate database interface module adapts each internal database schema to a uniform data interface used within Delta-X. Conversion routines are GIS-specific. Note that the conversion routines and the database schema are the only entities in Delta-X that are specific to geomatic data and GIS systems.

Note that the clear separation of data storage functions (the Delta-X server) from data processing functions (GIS) in a network not only facilitates data sharing and creation of logically organized application-specific data repositories, but it also enables us to take advantage of the latest development in both of the DBMS and GIS worlds. For example, the RDBMS currently used in the Delta-X prototype can be replaced with an OODBMS without any impact on the clients and other servers. In fact, different Delta-X servers using different database management systems can seamlessly co-exist in a Delta-X federation and exchange data with each other. Similarly, Delta-X can support a new type of GIS without the need to change the internal data interchange format and schema. These features allow different organizations with different data processing systems and/or architectures to join the same Delta-X federation and share data with each other.

2.3 The Communication Framework of Delta-X

The mode of communication between a Delta-X client and a Delta-X server or between Delta-X servers is asynchronous. Delta-X supports conversational transactions, which proceed as follows. A client connects to a server, requests a transaction, and then disconnects from the server. When the server finishes the requested transaction, or when it has to report the transaction's status, the server connects to the client and delivers requested data or reports the status.

The main reason for introducing the asynchronous communication paradigm into Delta-X is the long duration of the conversion/transfer transactions. A typical transaction can take many hours to complete and it would not be reasonable to maintain a connection between a client and a server and thus tie up network resources for the whole duration of the transaction.

The asynchronous communication between the server and the client has several other advantages. First, the server can simultaneously process multiple transactions from multiple clients without running out of communication channels. Second, asynchronous communication between Delta-X entities facilitates nesting and chaining of Delta-X transactions.

For example, a client may request a transaction from a certain server. The server starts the transaction and once in a while connects to the client to deliver a report on progress of the transaction. When the transaction migrates to another server, for example, to perform a special conversion, the new server will also connect to the client in the same way to deliver status reports and eventually the results. The client does not have to tear down the connection to the old server and establish a connection to the new server. However, the client must be informed by the servers that the transaction has migrated so that it knows where to send inquiries about the transaction.

The implementation of client-server and server-server asynchronous communication is based on the SUN remote procedure calls over TCP/IP.

3.0 The Delta-X Server

The Delta-X server is a node in a network of Delta-X hosts which has a data repository containing a part of shared data stored in the Delta-X federation. The Delta-X server performs the following functions:

1. Transaction management and surveillance
2. Data storage, access, and management
3. Data conversion
4. Client-server and server-server communication over a LAN or WAN
5. Naming and authentication for Delta-X clients

The Delta-X server comprises modules shown in Figure 2. The modules are: the Transaction Monitor, multiple Data Conversion modules, the Data Mover module, the Delta-X database management module, and one or more optional specialized GIS server modules to perform functions that are specific to that particular GIS. For example, a part of the conversion into ARC/INFO format must be performed on the ARC/INFO GIS system itself. The GIS server

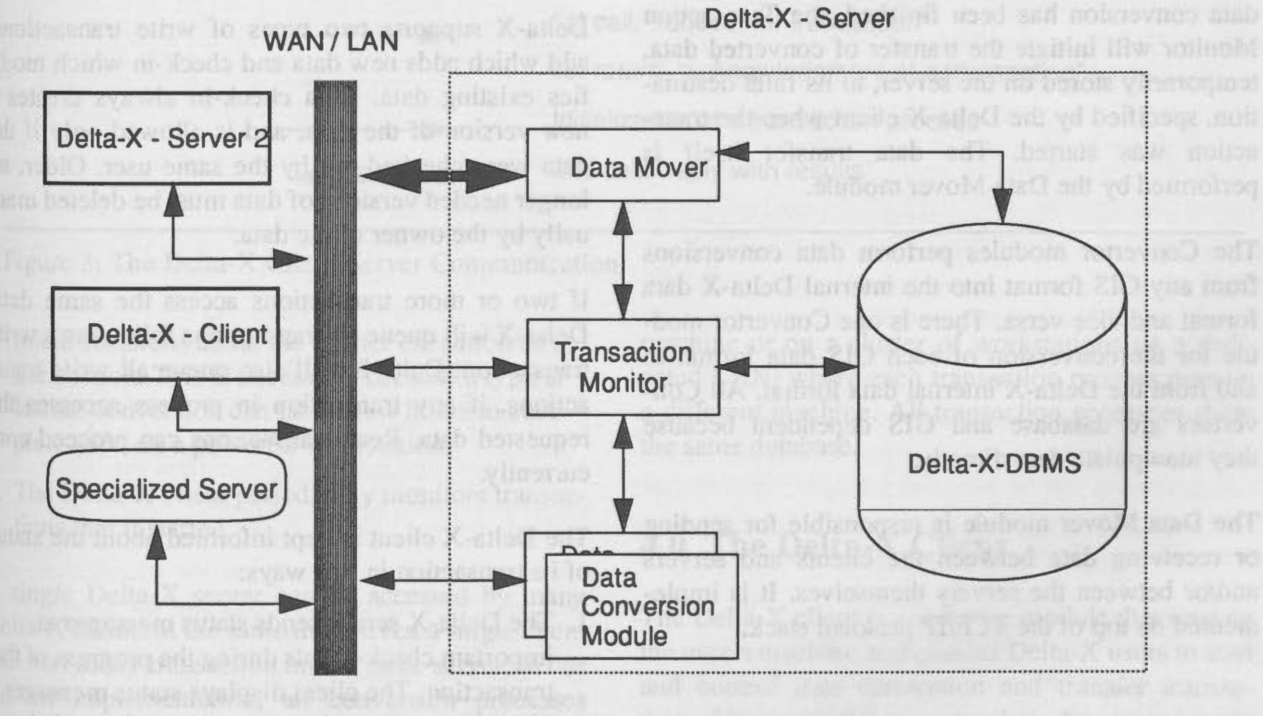


Figure 2: The Delta-X Server

modules can be physically located on the same machine as the other server modules, or they can run on specialized machines connected to the server through a LAN.

3.1 Delta-X Server Modules

The Transaction Monitor module performs transaction management tasks, such as transaction scheduling, transaction monitoring, authentication of transaction requests, locking of data (see Section 3.2), resource administration, and transaction commits, aborts and recoveries. The Transaction Monitor receives requests for new transactions from Delta-X clients or requests to process already running transactions from fellow servers in the Delta-X federation. The Delta-X server first validates the requestor's permissions for accessing the requested data and, if successful, it allocates resources for executing of the transaction and starts the transaction. The Transaction Monitor also notifies clients about the status of transactions in progress, finished transactions, or aborted transactions. The Transaction Monitor also logs accounting data in order to support billing of users who access data stored in the federation.

If required, the Transaction Monitor will relocate the transaction to another server, for example, to perform a conversion or finish a data transfer. After the data conversion has been finished, the Transaction Monitor will initiate the transfer of converted data, temporarily stored on the server, to its final destination, specified by the Delta-X client when the transaction was started. The data transfer itself is performed by the Data Mover module.

The Convertor modules perform data conversions from any GIS format into the internal Delta-X data format and vice versa. There is one Convertor module for the conversion of each GIS data format to and from the Delta-X internal data format. All Convertors are database and GIS dependent because they manipulate data directly.

The Data Mover module is responsible for sending or receiving data between the clients and servers and/or between the servers themselves. It is implemented on top of the TCP/IP protocol stack.

3.2 Delta-X Server Operation

The Delta-X server is started by when the operating system boots up. After starting up and initializing its communication and authentication tables, the server is ready to accept requests from clients.

Each Delta-X client calls a Delta-X server with a request to initiate a transaction (see Figure 3). The information sent in the request contains: Delta-X client's identification, description of the transaction, location and format of the source (in case of data import to Delta-X) or destination (in case of data export) data. After a successful authentication of the user which started the Delta-X client and of the type of the request, the Delta-X server sends an acknowledgment to the requesting Delta-X client and then starts processing the user's request (see Figure 3).

In the request, the data being accessed in the transaction may be locked, in which case the read and/or write access to the data (as specified in the lock request) is disabled for other users until the data is unlocked.

Delta-X supports three types of read transactions: normal read which only allows access to unlocked data, "dirty" read which allows read access to unlocked and write-locked data, and check-out read which can read and lock unlocked data.

Delta-X supports two types of write transactions: add which adds new data and check-in which modifies existing data. Data check-in always creates a new version of the data, and is allowed only if the data were checked-out by the same user. Older, no longer needed versions of data must be deleted manually by the owner of the data.

If two or more transactions access the same data, Delta-X will queue all transactions following a write transaction. Delta-X will also queue all write transactions, if any transaction in process accesses the requested data. Read transactions can proceed concurrently.

The Delta-X client is kept informed about the status of its transaction in two ways:

1. The Delta-X server sends status messages at important check-points during the progress of the transaction. The client displays status messages from the server so that the user is informed about the status of the transaction so that the user can

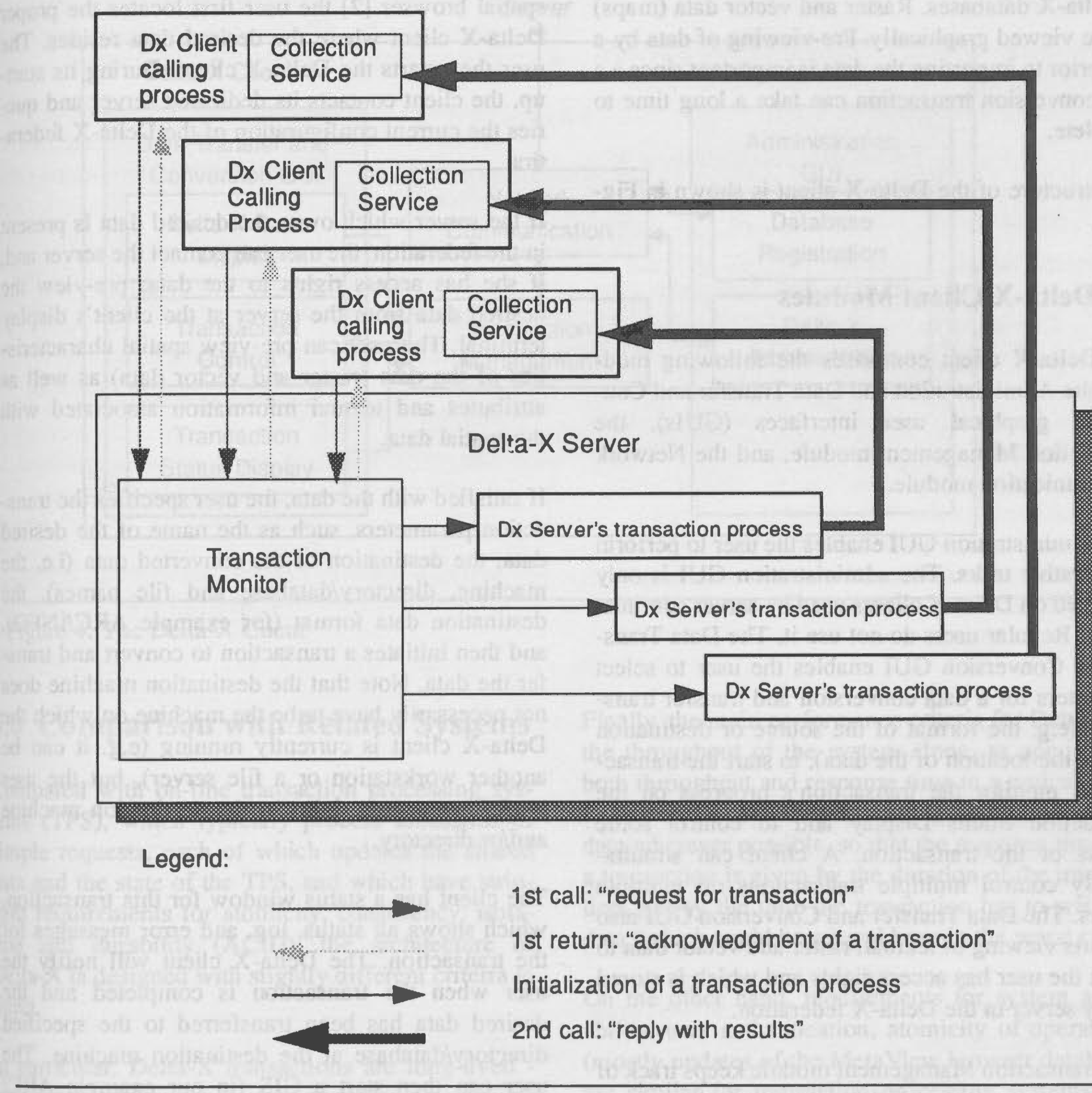


Figure 3: The Delta-X Client-Server Communication

make decisions about the further execution of a transaction. This is necessary, because a typical format conversion can take a few hours to complete even on a powerful workstation.

2. The Delta-X client periodically monitors transactions that it started.

A single Delta-X server can be accessed by many Delta-X clients at the same time. Even a single client can start many transaction on the same server. In our current implementation, all conversion processes share the same CPU. The Delta-X server architecture, however, can also be implemented on a parallel machine or on a cluster of workstations on a dedicated LAN, where each transaction process runs on a different machine. All transaction processes share the same database.

4.0 The Delta-X Client

The Delta-X client is a software module that runs on the user's machine and enables Delta-X users to start and control data conversion and transfer transactions. Users can either export their data to or import data from a server. The Delta-X client also supports administrative functions, such as joining or leaving the federation or authorizing user's access to data stored in the federation, and viewing of data stored in Delta-X databases. Raster and vector data (maps) can be viewed graphically. Pre-viewing of data by a user prior to importing the data is important since a a data conversion transaction can take a long time to complete.

The structure of the Delta-X client is shown in Figure 4.

4.1 Delta-X Client Modules

The Delta-X client comprises the following modules: the Administration and Data Transfer and Conversion graphical user interfaces (GUIs), the Transaction Management module, and the Network Communication module.

The Administration GUI enables the user to perform admistrative tasks. The administration GUI is only packaged on Delta-X clients used by server administrators. Regular users do not use it. The Data Transfer and Conversion GUI enables the user to select parameters for a data conversion and transfer transaction (e.g. the format of the source or destination data or the location of the data), to start the transaction, to monitor the transaction's progress on the Transaction Status Display and to control some aspects of the transaction. A client can simultaneously control multiple transactions on multiple servers. The Data Transfer and Conversion GUI also supports viewing of textual, raster and vector data to which the user has access rights and which is stored on any server in the Delta-X federation.

The Transaction Management module keeps track of all active transactions. If the client is terminated and there are active transactions started by the client, the Transaction Management module will cancel all outstanding active transactions.

The network communication module performs all communication tasks with servers. The communication functions can be either requests to view spatial data stored in one of the servers, or requests to perform data transfer/conversion transactions.

4.2 Delta-X Client Operation

The mode of operation for the Delta-X client is as follows. Suppose a user wants to import some GIS data, stored on one of the Delta-X federation servers, into her GIS. Suppose the GIS at the user's workstation is ARC/INFO. Using the MetaView spatial browser [2] the user first locates the proper Delta-X client where the desired data resides. The user then starts the Delta-X client. During its startup, the client contacts its dedicated server and queries the current configuration of the Delta-X federation.

If the server which owns the desired data is present in the federation, the user can contact the server and, if she has access rights to the data, pre-view the desired data from the server at the client's display terminal. The user can pre-view spatial characteristics of the data (raster and vector data) as well as attributes and textual information associated with the spatial data.

If satisfied with the data, the user specifies the transaction parameters, such as the name of the desired data, the destination of the converted data (i.e. the machine, directory/database and file names), the destination data format (for example ARC/INFO), and then initiates a transaction to convert and transfer the data. Note that the destination machine does not necessarily have to be the machine on which the Delta-X client is currently running (e.g. it can be another workstation or a file server), but the user must have access rights to the destination machine and/or directory.

The client has a status window for this transaction, which shows all status, log, and error messages for the transaction. The Delta-X client will notify the user when the transaction is completed and the desired data has been transferred to the specified directory/database at the destination machine. The user can then start a GIS (in our example ARC/INFO) to load the converted data into the GIS.

Similarly, when a user wishes to export data to the Delta federation, she first consults the MetaView spatial browser to determine where the data should be located. If the user has access and update rights for the specified server, she specifies the transaction parameters, such as the source data format, and location (the machine, directory and file names), the destination format and starts the export transaction. The Delta-X client has a status window for this transaction, which shows the progress of the transaction. The Delta-X client will notify the user when the transaction is completed and the desired data has been stored on the specified server.

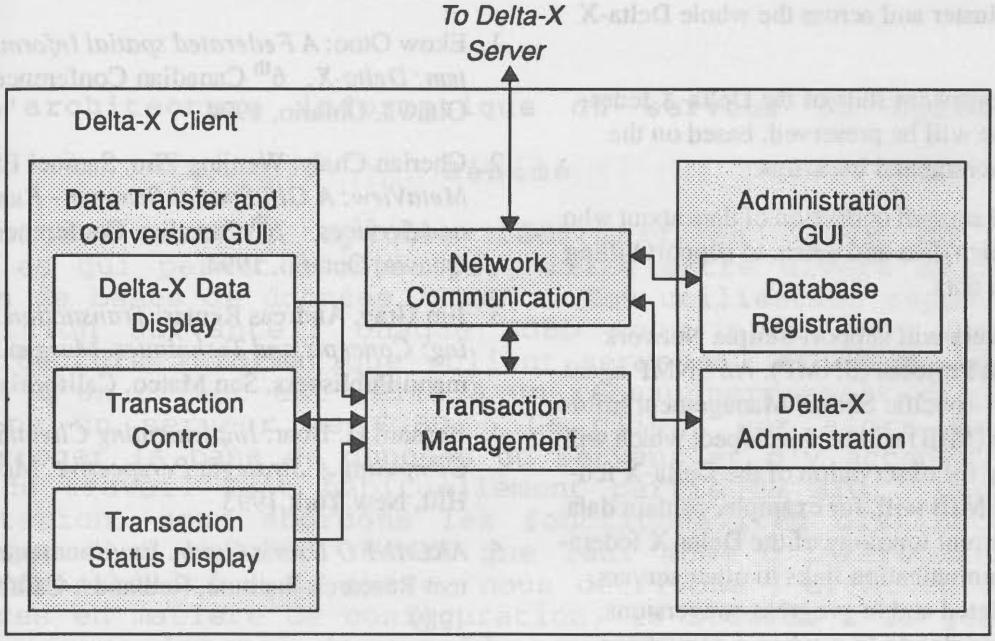


Figure 4: The Delta-X Client

5.0 Comparison with Related Systems

Compared with on-line transaction processing systems (TPS), which typically process thousands of simple requests, each of which updates the shared data and the state of the TPS, and which have stringent requirements for atomicity, consistency, isolation and durability (ACID), the architecture of Delta-X is designed with slightly different criteria in mind.

In particular, Delta-X transactions are long-lived - they can take about an hour to complete on a powerful workstation. Therefore, mechanisms must be put in place for the user to inquire about the status of the transaction and to control the flow of the transaction. Moreover, if possible, Delta-X must allow concurrent accesses to data for as many transactions as possible so that reasonable throughput and response time of the system is guaranteed. This requirement can be easily met, as most accesses to Delta-X data are read-only.

Second, Delta-X supports versioning of data - each update creates a new version of the data. Thus, data consistency problems are avoided.

Finally, the main performance criteria for Delta-X is the throughput of the system alone, as opposed to both throughput and response time in a typical TPS. However, Delta-X avoids waiting and queueing for data wherever possible, so that the response time for a transaction is given by the duration of the transaction, not by the time the transaction has to wait for data (which could be several hour in the worst case).

On the other hand, requirements for system availability, user authentication, atomicity of operations (mostly updates of the MetaView browser database) are similar for transaction processing systems and Delta-X.

6.0 Future Work

The prototype Delta-X implementation will be gradually extended in the following areas:

1. The support for high security will be added to make Delta-X C2-compliant.

2. Delta-X servers will be made fault-tolerant and highly available by adding hot stand-by servers and/or introducing data replication across the servers in the federation.

3. Support for parallelism and load-balancing within the cluster and across the whole Delta-X federation.

4. The latest consistent state of the Delta-X federated database will be preserved, based on the automatic incremental back-ups.

5. Delta-X will support collection of data about who accessed which data and when, to support billing for access to data.

6. Delta-X servers will support Simple Network Management Protocol (SNMP). An SNMP agent and a vendor-specific SNMP Management Information Base (MIB) will be developed, which will serve mainly for observation of the Delta-X federation. The MIB will, for example, contain data about the current topology of the Delta-X federation, the communication links to other servers, about completed and in-progress conversions, software error alarms (e.g. when a conversion routine crashes), data error alarms (e.g. when a conversion routine encounters a data inconsistency), billing data, etc.

7.0 Conclusion

The paper describes a flexible and powerful client/server architecture which is targeted for two groups of users: distributors of geographical data and their customers (users). The distributors get a powerful automated data distribution system. Their customers can electronically download data in a format that is most suitable for them. The Delta-X architecture is flexible, and can be easily extended to support other types of applications.

Acknowledgments

We would like to thank Ekow Otoo for his inspiring ideas and comments which he always willing to share. We are also thankful to GIS Division, Energy and Mining Resources Canada for financial support of this project and last but not least the thoughts of gratitude lead to our wives Ivana and Maria for their understanding for our work.

References

1. Ekow Otoo: *A Federated spatial Information system: Delta-X*, 6th Canadian Conference on GIS, Ottawa, Ontario, 1994

2. Cherian Chaly, Wenjing Zhu, Samuel Effah: *MetaView: A GIS Spatial Browser - Functions and Services*, 6th Canadian Conference on GIS, Ottawa, Ontario, 1994

3. Jim Gray, Andreas Reuter: *Transaction Processing: Concepts and Techniques*, Morgan Kaufmann Publishers, San Mateo, California, 1993

4. Bernard H. Boar: *Implementing Client/Server Computing.A Strategic Perspective*. McGraw Hill, New York 1993

5. *ARC/INFO User's Guide*. Environmental System Research Institute, Redlands, California, 1989

6. *CARIS Command Reference*. Universal Systems, Fredericton, New Brunswick, 1992

L'architecture informatique du serveur du Delta-X

Résumé

Delta-X est un système fédératif de bases de données multiples qui permet l'interopérabilité entre divers systèmes de gestion de bases de données (SGBD). Son utilisation repose sur la supposition suivante : chaque SGBD sous-jacent est intégré à l'architecture informatique «client-serveur», et chaque poste de travail d'un client est relié à un réseau permettant d'accéder à au moins un serveur de SGBD. Le client est alors en mesure d'interroger la base de données du serveur et d'y accéder, et son poste de travail fait habituellement partie du SIG. Dans cette présentation, nous abordons les fonctionnalités distinctes d'un client et d'un serveur, ainsi que leur mode d'interconnexion et d'interopérabilité. En outre, nous décrivons l'architecture, les exigences en matière de configuration, le soutien à la gestion de réseau (comme l'appel de procédures à distance) et les services rendus d'un serveur à un autre, d'un client à un serveur et d'un client à un autre.

Mots-clés : bases de données réparties, interopérabilité, architecture «client-serveur», appel de procédures à distance, service de répertoire mondial.

MetaView/GIS: The GIS Spatial Browser

Its Design, Implementation and Operation

Samuel Effah,
Cherian Chaly,
Wenjing Zhu

Geographic Information System's Division
Natural Resources Canada,
615 Booth Street,
Ottawa, Ontario, K1A 0A9

Abstract

Facility for browsing metadata helps users to save their valuable time identifying and locating data they need. MetaView/GIS is a spatial browser that provides this facility to users. It provides access to metadata of various databases, namely, information on specific data sets, location, ownership, format, cost of access, etc. A user can query the metadata to access further information not directly visible via the user interface. Visualization of data before exchange is also important to most users in order to decide whether the data is suitable or not. This can be done by displaying it through a complementary software. Although the computing, data communications and graphics technologies provide the opportunity to do just that, not many organizations have resolved to achieve this. GIS Division provides a solution to these and similar questions with an implementation of a GIS Spatial Browser which serves as a front end to a network of spatial topological databases. This paper will discuss the philosophy behind the design of MetaView/GIS, its functionalities and services.

Keywords: Metadata, Spatial, browser, database, data set, GUI

1.0 Introduction

Geographic Information System (GIS) is a fast developing technology with unbounded area of application areas. One of the major problems confronting this technology is in the area of identifying the source of data relevant for the development of an applications.

There is therefore the need for a spatial browser, that will provide GIS application developers, the means of searching and locating data from various sources. MetaView/GIS is a software that provides users, a friendly interface for this purpose. It reduces drastically the time and effort needed to seek desired information on spatial data sets. MetaView/GIS does not disclose the contents of a database to allow retrieval of a data set. It only provides access to the metadata of the data set. Such data may include the following:

- the format and scale of available map data,
- the name of the database and site information of the data set,
- the agency/organization and contact persons of the data set,
- the projects and applications that have utilized the data sets,
- the documents and guides on the data set,
- the source and coverage of the data set,
- the access policy, protocol and cost of data set,
- etc.

MetaView/GIS is configured as a client-server model to run over Internet. The client is a user Unix workstation running a client component. The server component runs on MetaView/GIS site. A number of other client-server directory services exist for locating and retrieving information across the Internet. Such directory services include Wide Area Information Server (WAIS)[5], World-Wide Web (WWW)[4], Archie[7] and Gopher[6]. WAIS, Archie and Gopher provide the user with an overview of likely places to find the desired information, and then help the user locate the specified information items. In WWW, a GUI software known as *xmosaic* allows a user to navigate through databases with a mouse click in a hypertext mode.

MetaView/GIS differs from these systems in a number of ways. First, MetaView/GIS, operationally, serves as a partial front end of a commercial service DeltaX[1,2]. Second, it provides an X11 graphical interface at the user's client which relies on commercial databases management system and not a hypertext system. It conforms to an information structured standard specified by the Canadian Inter Agency Committee on Geomatic (IACG). Third, users can specify spatial queries. Fourth, MetaView/GIS provides more information related to a data set than those existing in these directory services. However in designing MetaView/GIS, the services provided by these systems were taken into consideration and in fact are utilized by the MetaView/GIS system. For example, if a data set is available free at a site, one may choose to retrieve it by employing the services of the other directory services.

In this paper, we present some of the functionalities, design concepts and implementation of MetaView/GIS. Section 2 discusses the design concepts, the intended users of MetaView/GIS. User's interaction issues are addressed in section 3. Section 4 deals with administrative aspects of the MetaView/GIS. Section 5 summarizes this paper and presents some projected enhancement.

2.0 Operational concepts MetaView/GIS

One of the main difference between MetaView/GIS and the other directory services is the ability for a user to specify queries using region selection. MetaView/GIS maintain two types of databases: one contains map data, the other contains the metadata. The map database contains data for displaying maps around the world, including maps of the continents, regions and the globe. The other database maintains the directory, catalogue, etc., of data sets in other databases. This description includes information such as who owns the data set, how it is organized, how much it will cost to access and how it can be accessed, etc. The first database is internal to the MetaView/GIS system, and users cannot therefore influence its organization and/or modification. The metadata database is however available for updates by certain category of users.

One unique characteristics of MetaView/GIS, is that, beside the graphical interface of map display, region selection, point-and-click selection using a mouse device and a menu selection, advanced users can utilize the full power of SQL to specify customized queries. There are two main users of the MetaView/GIS system. Those who will maintain the data in the database and those who will interact with the database to access information about data sets and databases.

The MetaView/GIS software is organized on a client/server model. The server runs at a site having typically a huge volume of data to be sold or made publicly available. The server has the MetaView/GIS databases. A client (which is free to access information) runs at the user's site. The user invokes the MetaView/GIS system by issuing the command *xmvgis*. Queries by the users are bundled and transferred to the server, which retrieve the necessary data from the database and returned to the client. The communication between the client and the server is done in an asynchronous mode over a communication network. Figure 1 shows how the system is modeled. The figure shows a network of MetaView/GIS clients and servers in a Local Area and Wide Area Networks.

3.0 User Interaction

3.1 Server

The MetaView/GIS server interfaces between the MetaView/GIS clients and the databases of the system. Basically, the server is responsible for:

- transaction management; this includes the monitoring and scheduling of transactions, the authentication of client requests, the administration of resources and the commit/abort of transactions and recovery.
- store, request executions and data management.
- connecting to a client to transmit data and communicating with other MetaView/GIS servers for assistance when necessary.
- authentication of MetaView/GIS privileged clients.

3.2 Client

The client includes the GUI portion of the MetaView/GIS software that runs at the user's site. Where a client's workstation cannot support X11/Graphics display, a user invokes MetaView/GIS by the command *mvgis*. When MetaView/GIS is started it first displays its top level window which is shown in figure 2. At this point the client is attached to the defaults MetaView/GIS server. At the top level, a user has the option of connecting to DeltaX, a complimentary system of MetaView/GIS. DeltaX provides interoperability among Geographical Information Systems by converting data stored in one format and transporting data to users. After user authentication, a user can perform some system administrative functions by clicking the *System* button at the top level window. To exit the MetaView/GIS system, a user must select the *Quit* button in this window. We will discuss the *Help* button shortly. To start the MetaView/GIS itself, one must select the *Meta-Data* button at this top window. This button pops up the Database Information Window shown in figure 3.

3.3 Queries

When the Database Information Window is displayed, a default server is connected to the client. The user can select a desired server by clicking on the *Servers* button in the window. All queries in the session will be sent to the connected server. To search the chosen server's database, a search option must be specified. One can search the database by any of the following options: *keywords, data set name, category code, category name, agency/owner name, contact person name, database name, application name* or *project name*. The *Summary* button gives a summary of the data sets in the database. Apart from querying the database for the summary of the data sets, for each search option specified, the corresponding entry must be entered in the entry field that follows the option button. This field is marked "*Enter options:*" in figure 3. Entries in this field need not match exactly the content of the database. A partial match is sufficient for a successful search. The search string entered can be in any case, upper or lower case, since the field is case-insensitive. On a return (pressing the Enter key) or a click on *Accept* button, the search will begin. While an operation is in progress, e.g., a search of the database, the user may proceed to perform other operations.

The results if any, is displayed in the scrollable window marked "*Search Results:*" in figure 3, otherwise a message of no match is displayed. The results returned are the data sets that qualify given the search criteria. The results for any search, except for the *Summary* option, are a list of qualifying data set keys (as we will discuss later, these keys may be useful to users with SQL knowledge), acronyms and data set names. The *Summary* search result is a list of category code, category name and a count of data sets in each category. Having received the search results, the user may proceed to view or retrieve further information on the data set that interest him/her. These include information about the owner of the data set, the contact persons of the data set, available documents, applications and projects of the data set and the data sets storage and access informations.

To access the information, the user must highlight the data set that he/she is interested in (at the scrollable list window) and then select the desired information at the area marked "*Retrieved info:*" in figure 3. The *All* button will automatically select all the available information for the user. After selecting the information desired, a click on the *Retrieve* button will send the request to the server. If the desired information is available, it is displayed in one or a number of windows in turn at the user's workstation.

The sequence of actions to be followed at a client workstation to retrieve information is illustrated as follows. Assuming that the user has invoked the *Database Information* window shown in figure 3. Let the *Search option* chosen be the *Dataset Name*. If the string entered is "canad", then the display in the *Search result* of the window shows the information returned as the result of the query. If the user now highlights the line shown, i.e., "*Canadian Vessel Marine Surface* ...", then further information about this data set may be retrieved. Suppose the user now clicks on the buttons *Owner* and *Contact* and then request that information be retrieved, the result of such a query is the display of the two windows in Figures 4 and 5 respectively. The information in these windows correspond to contents of forms that can be printed.

When the required information windows are displayed, the user has three button for the next sequence of operation. The user may choose to close the window, get detailed help about the windows, or save the contents of the window in his/her working directory. If a user elects to save the information, the save file corresponding to the window is given as *<window-name>.mv*, where <window-name> is the information being viewed. For example, if we save the owner information of figure 4, the file name will be *owner.mv*. Any subsequent saving of owner information of any other data sets will be appended to this file. Detailed information on the data set itself is provided in three windows. Therefore, if detailed information on a data set is requested (i.e., selecting *Dataset* at *"Retrieve info:"* in figure 3), one window only is first displayed. The user can access the other information by selecting one of the two buttons labeled *Other Info* and *Storage Info*.

If the information say, about the owner or contact person (see figures 4 and 5) is displayed, and a new data set is selected from the list in figure 3, requesting *Retrieve* with *Contact* and *Owner* still enabled, will cause the already displayed windows to be updated with the information on the new data set.

3.4 Visual Queries

Visual queries involve first invoking a display of a globe, and then determining a region or an area of a globe where information is desired. A user can invoke this facility of MetaView/GIS by opening the map database. The button marked *"Spatial Browser"* pops up the map window. In this window a user can display any region of choice. Using a mouse and a rubber banding selection technique, a user may zoom in on an area to facilitate more precise region selection. The selected area is interpreted as the bounding coordinates and is transmitted to the *"Search region:"* in the Database Information Window. A user can also specify the search region manually for searches to be constraint within the region. The provision of visual spatial querying capability is one of the main distinguishing features of MetaView/GIS.

3.5 Getting Help

Each window of the MetaView/GIS interface has a help button. In addition most buttons, in display text areas or labels have individual help messages associated with them. These messages appear when the buttons are mapped (i.e., when the cursor enters the boundary of the button). These help messages appear at the help message string in the upper part of most windows. For example, in figure 6, the text *"Scroll the text information upwards"* is the message string of the *Pg-Up* button. The message will be undisplayed if the cursor moves away from the *Pg-Up* button. MetaView/GIS help is organized in an index/contents format. When the help button of a window is clicked, the help window pops up, with the required window's index entry and contents shown as the default. While the help window is displayed, a user can assess other help information by selecting the appropriate index entry. The help window is shown in figure 6 with MetaView/GIS mailtool's help contents displayed.

3.6 Update Functions

There are essentially, two classes of users of MetaView/GIS. A general public user and a privileged user. The privileged clients are the database owners that provide the metadata of their databases/data sets to the MetaView/GIS database. These database vendors (i.e., owners of databases or the data sets) need to insert data in MetaView/GIS database and also update existing information. MetaView/GIS provide a more flexible approach and friendly entry forms for these purposes. To insert a new data or modify an existing information, such privileged users must pass a password validation check. All database vendors have password that allows them to gain access to the MetaView/GIS update functions.

To modify information (say a contact person's data) on a particular data set, that data set must be selected from the *"Search Result:"* list in the *Database Information* window. Selecting of the *Update/Modify* option from the pull-down menu pops up the *Update* windows Header window as shown in figure 7. The user must then enter a valid password and then select the data to be modified. For example, if the user clicks on the *Person* button, this pops up a window with the current data values corresponding to the concerned data set. Changes can now be made by the user. The modified data or newly inserted data is logged in temporary tables in the database. These changes are actually effected by the database administrator (DBA).

3.7 Other MetaView/GIS facilities

The main window of the browser, allows users to make request by using either and on-line or off-line tool. A user can *ftp* or *telnet* to other machine to transfer data. These systems use third party networking software for their functions. An off-line request can also be made by a FAX. The system prepares a fax report for the user based on the data set that is presently being accessed. A handy mail-tool is also available for users. These functions can be accessed with the menu *Request/Off-line/Fax* and *Request/Off-line/Mail*.

The browser also has gateways to other directory services. Briefly these services are:

- SQL: This is provided for advance users who may wish to interact with the MetaView/GIS databases directly. Based on the server's vendor database.
- DeltaX: A tool for the interchange of GIS data formats. A complementary system of MetaView/GIS.
- GCNet: A directory service developed at SMRSS, NRCan for researchers, scientists and users of remote sensing information.
- Mosaic: An Internet information browser and WWW client of NCSA. A user can connect to most directory services from here.
- Archie: The Archie information system is a network-based information tool offering pro-active data retrieval and indexing for widely distributed collections of data. Perhaps the best known application of the Archie system is to maintain the Internet Archives database.
- FreeNet: A National Capital FreeNet information service based in Ottawa, Canada. Users can connect to other FreeNets across the globe.

4.0 Administrative functions

The administrative functions in MetaView/GIS include the registration of new privilege users and the maintenance of the MetaView/GIS databases. A privilege database vendor must be registered and given an account so as to enable him/her to input his data into the database. The administrator also runs programs that will reflect modification of existing data items and changes to the database.

Apart from the DBA, no other user of the system can issue a commit transaction on the spatial data that is assessed in the database. Transactions of the database vendors commit, into tables that are not assessed by other users of the system. All tables are created with journaling option. Thus the underlying database management system (Ingres and Oracle) keeps a record of all changes to the tables. The DBA uses these journals for monitoring updates and for maintaining change histories.

A strict set of referential integrity rules are defined on the database tables. As much as the system can, new data items are cleaned up. On failure to meet the integrity rules, a message is transmitted back to the database user about the problem. Since the DBA (of the GIS vendor site) is the only one who can update the MetaView/GIS databases, the autonomy of the database is always maintained. The disadvantage here is that changes made by the database vendors do not get reflected in the MetaView/GIS database immediately.

5.0 Conclusion

This paper has described a new spatial browser and has presented its functionalities. The MetaView/GIS directory service provides a flexible GUI for querying Geographical Information Databases by users and for updates by database vendors. Unlike other directory services, a user of MetaView/GIS can query the databases within a constraint region by manipulation of the visual map display.

Planned enhancement to the current implementation include:

1. extension of MetaView/GIS database to include free text data and to include function for handling document retrieval.
2. use of third party software to expand the capability of the server. This will see MetaView/GIS tolerating other database other than what is currently supported under relational DBMS.
3. use of third party software to convert our present help facility to hypertext based.
4. porting of various routines to non-GUI, to enable other systems without X terminals to access MetaView/GIS databases.
5. the implementation of a client to run under WindowNT and Windows.

Acknowledgment

Our sincere thanks goes to Ekow Otoo and the team members of both the DeltaX and MetaView/GIS projects. We will also like to thank the GIS Division of Natural Resources Canada for their support for these projects.

References

1. Otoo, E. *"A Federated spatial Information system: DeltaX"*, a paper to be presented at the 6th Canadian Conference on GIS, Ottawa, Ontario, 1994
2. Medved, J. and Petras, J. *"The client-server architecture of DeltaX"*, a paper to be presented at the 6th Canadian Conference on GIS, Ottawa, Ontario, 1994
3. Sharon Fisher. *"Riding the Internet Highway: A complete guide to 21st Century Communication"*, New Riders Publishing, 1993.
4. T. Berners-Lee, R. Cailliau, J.-F. Groff, and B. Pollermann, *"World-wide web: the information universe"*, Electronic Networking: Research, Applications and Policy, vol. 2, pp. 52--58, Apr. 1992.
5. C. Stanfill, *"Massively parallel information retrieval for wide area information servers"*, in Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, (Charlottesville, Virginia), IEEE, Oct. 1991.
6. F. Anklesaria, M. McCahill, P. Lindner, D. Johnson, D. John, D. Torrey, and B. Alberti, *"The internet gopher protocol (a distributed document search and retrieval protocol)"*, Request for Comments (Informational) RFC 1436, Internet Engineering Task Force, Mar. 1993.
7. M. F. Schwartz, A. Emtage, B. Kahle, and C. B. Neuman, *"A comparison of Internet resource discovery approaches"*, Computing Systems (The Journal of the USENIX Association), vol. 5, pp. 461--493, Fall 1992.
8. *Ingres Reference Manuals.* Ingres Corporation, Alameda, California, 1992.

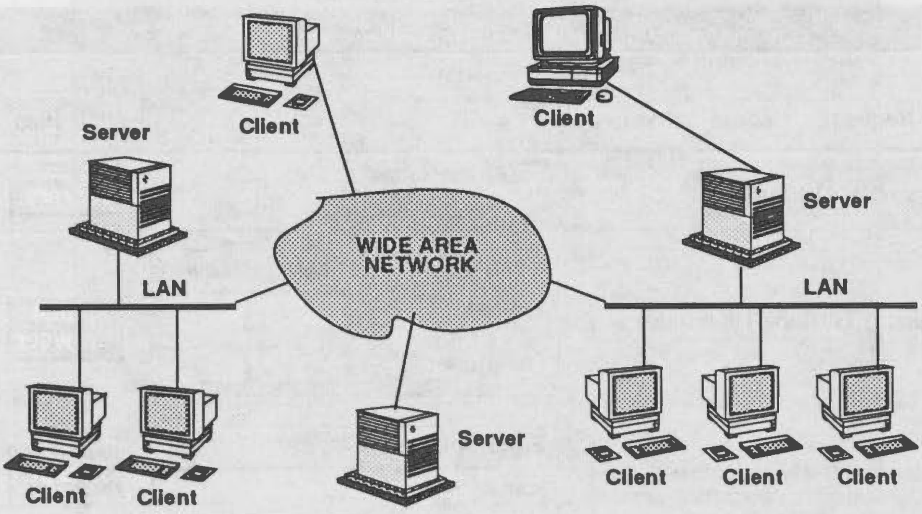


Figure 1: The client-server model of MetaView/GIS.

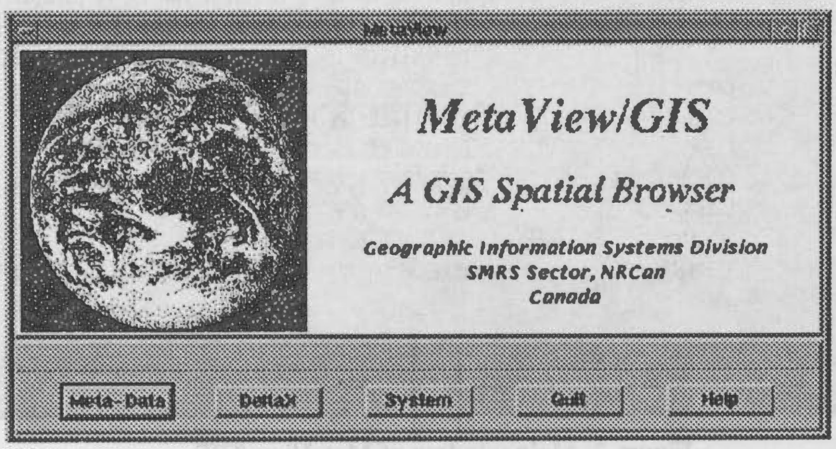


Figure 2: Top level window of MetaView/GIS

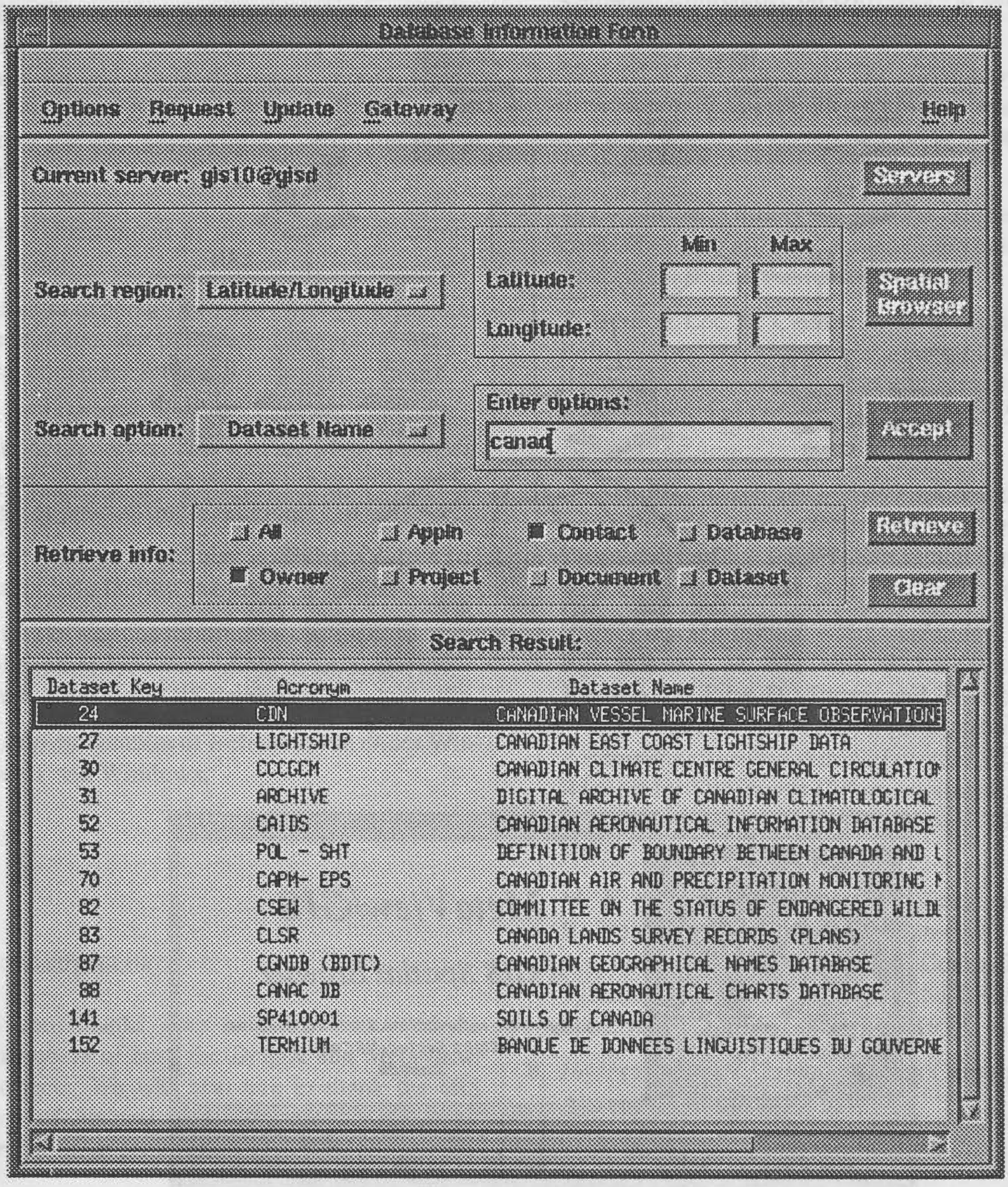


Figure 3: Main window of MetaView/GIS.

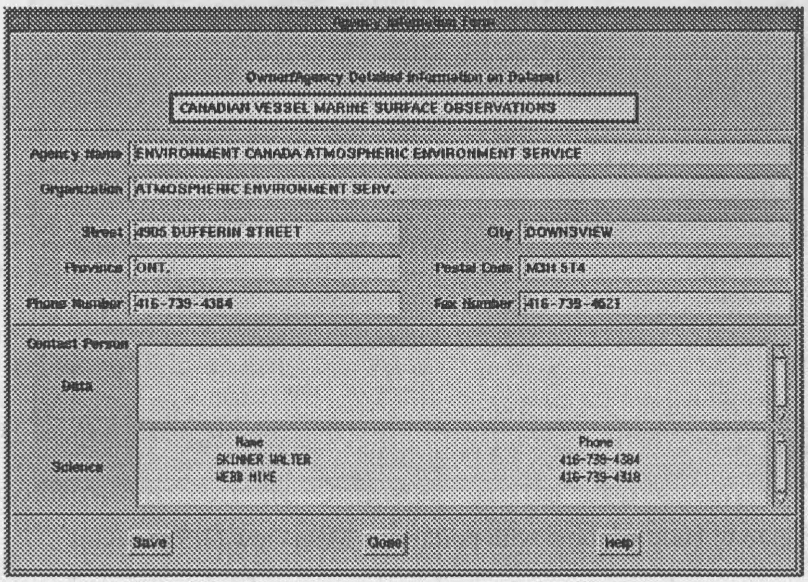


Figure 4: Agency retrieval window with data

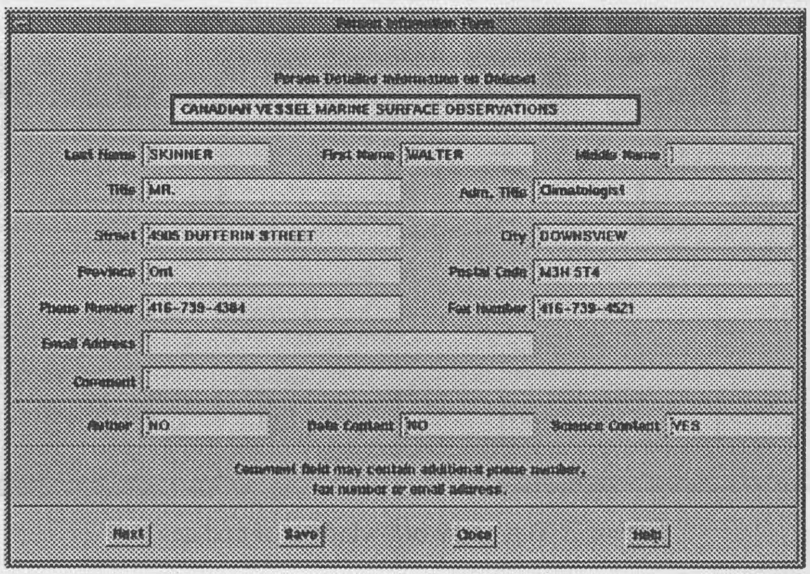


Figure 5: Person retrieval window with data

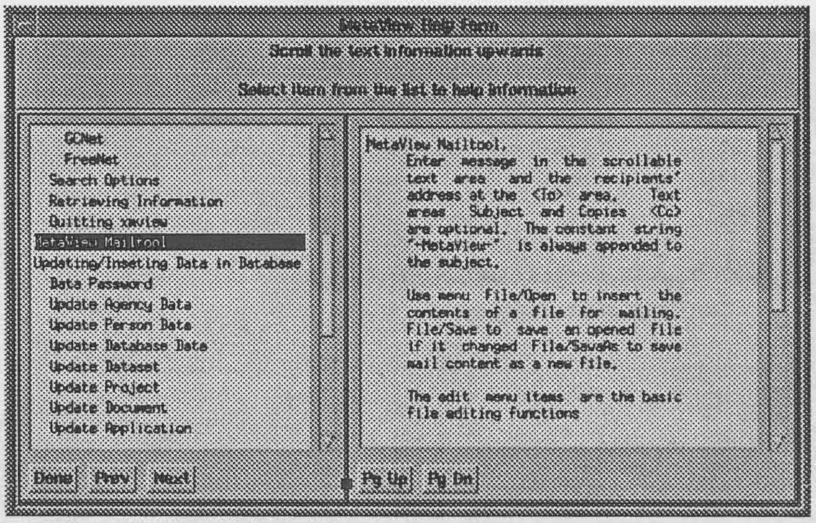


Figure 6: Help window with pointer on MetaView/GIS mailtool

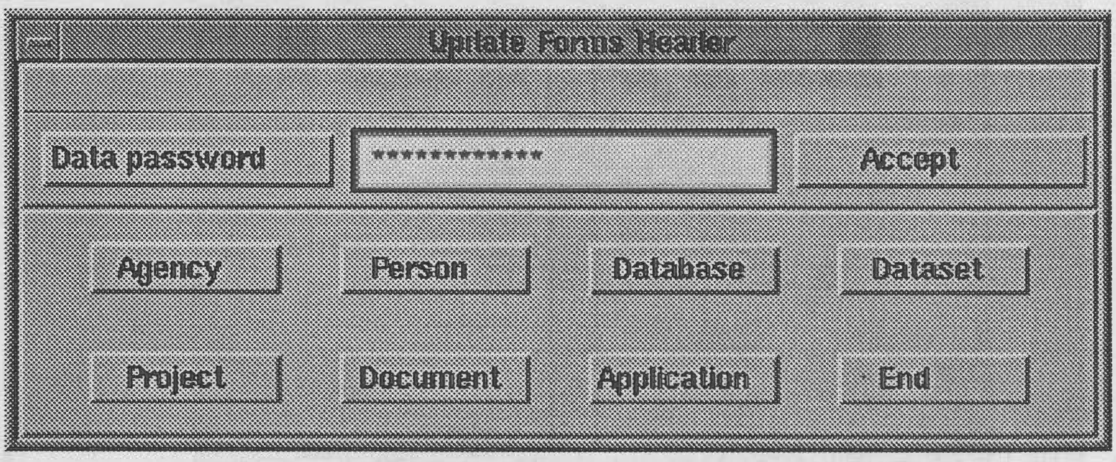


Figure 7: Top/password window of updates

Méta-vision : Un navigateur de données spatiales pour SIG - Fonctions et services

RÉSUMÉ

Une installation de survol de métadonnées économise le précieux temps des utilisateurs lors de l'identification et de la localisation des données dont ils ont besoin. Le Méta-vision (Metaview) est un navigateur de données spatiales offrant cette possibilité aux utilisateurs. Il donne accès aux métadonnées de diverses bases de données en fournissant de l'information concernant des ensembles de données spécifiques, dont l'endroit où ils se trouvent, le propriétaire, le format, les coûts d'accès, etc. Un utilisateur peut interroger les métadonnées pour accéder à de l'information non directement disponible par l'entremise de l'interface pour utilisateurs. La visualisation des données avant leur échange est également importante car elle permet à la plupart des utilisateurs de décider si les données leur conviennent; l'affichage des données le permet. Bien que les technologies de calcul, de communication de données et de représentation graphique permettent de ce faire, peu d'organismes y ont consacré des efforts. La Division des SIG a été en mesure de solutionner ces problèmes et des problèmes similaires par la conception d'un navigateur de données spatiales pour SIG qu'elle a placé au début de sa base de données topologique spatiale. Dans la présente communication la philosophie qui sous-tend la conception du Méta-vision est abordée et l'on discute les fonctionnalités et les services de ce système.

Mots clés : métadonnées, spatial, navigateur, base de données.

INTEGRATED GEOGRAPHIC INFORMATION SYSTEMS (IGIS): STATUS AND RESEARCH ISSUES

Manfred Ehlers
David R. Steiner
ISPA, University of Osnabrück-Vechta
Vechta, Germany
and
Nickolas Faust
Georgia Institute of Technology
Georgia Technical Research Institute
Atlanta, GA, USA

Commission II, Working Group 2

KEY WORDS: Geoinformatics systems, Integrated Geographic Information Systems (IGIS), Technology Advances

ABSTRACT

This paper presents status and trends for recent developments of integrated geographic information systems. Progress in computer hardware and software are the basis for a structural change in the systems aspect. Boundaries between systems for image analysis, geographic information processing, softcopy photogrammetry, and digital cartography are ceasing to exist. We will present challenges and possibilities for ***Geoinformatics Systems***.

KURZFASSUNG

Dieser Beitrag zeigt den gegenwärtigen Zustand sowie zukünftige Trends für die jüngeren Entwicklungen im Bereich integrierter geographischer Infortmationssysteme auf. Fortschritte im Hardware- und Software-Bereich sind die Basis für einen strukturellen Wandel: Grenzen zwischen Systemen für Bildverarbeitung, für geographische Informationsverarbeitung, für digitale Photogrammetrie sowie für digitale Kartographie beginnen zu verwischen. Wir werden Herausforderungen an und Möglichkeiten von ***Geoinformatics Systems*** darstellen.

1.0 WORKING GROUP ACTIVITIES

The activities of the ISPRS Working Group II/2 *'Hardware and Software Aspects of Geographic Information Systems'* were primarily directed toward the identification of significant trends for future developments and the integration with other systems for spatial data handling. Specific topics of interest included:

- Design and operational aspects of the integration of GIS with image analysis systems;
- Technologies for three-dimensional GIS;
- Studies of man-machine interactions: display techniques, interactive techniques, and audio interaction;
- Studies of benchmark design for integrated GIS;
- Studies of GIS characterized by workstation, mainframe, and microsystems in a heterogeneous environment;
- Studies of parallel processors, array processors, supercomputers, and optical hybrid systems for improving GIS;
- Hardware and software for input/output aspects of GIS.

Within this context, WG II/2 held three business meetings and organized an international workshop on *'Requirements for Integrated Geographic Information Systems'* which was held in conjunction with the 'Second Thematic Conference on Remote Sensing for Marine and Coastal Environments'. Fifty-eight participants from thirteen different countries attended the workshop which was held at the New Orleans Marriott Hotel on February 2 and 3, 1994.

Twelve oral presentations were given on a variety of topics related to the issues and challenges faced by researchers using integrated GIS for work in Belgium, Egypt, England, Germany, Hungary, Italy, Norway, and the United States. These topics included functional requirements for implementing and using GIS; the integration and use of various tools and techniques such as virtual reality, neural networks, and numerical modelling; metadata management; the use of temporal components; commentaries on the status and challenges of implementing GIS in several foreign countries; and case studies in wetland and aquaculture applications.

Vendors from six different software corporations also attended to participate in a panel discussion, which was a highlight of the workshop. An interactive poster session was also held, at which seven posters were presented on diverse areas of application. Proceedings for the workshop will be published later this year (Ehlers, et al., 1994).

Another workshop on 'Visualization and GIS' is tentatively scheduled for May of 1995 in Germany.

2.0 SYSTEMS TECHNOLOGY ADVANCES

A critical review of the systems aspect for the processing of geographic data reveals that progress in this area has been primarily driven by advances in technology. These advances come from fields remote sensing, photogrammetry, surveying, or mapping.

In particular, progress in computer science and the major acceptance of geographic information systems (GIS) as a unifying technology are challenging the "separation of disciplines". Consequently, the shift toward integrated systems for processing of geoinformation is driven by advances outside the disciplines traditionally represented by ISPRS. It is, however, imperative that ISPRS responds to these technological challenges.

We need an integrated approach for research, development, and education in geoinformation processing which might be coined "geoinformatics". If we do not aggressively adopt such an integration, photogrammetry and remote sensing might end up as minor subareas within computer science and informatics.

In the following sections we will try to summarize advances in hardware and software which generally apply to all systems involved in geoinformation processing.

2.1 Hardware Advances

2.1.1 Computational Power. The advances in speed and power of computers have been phenomenal. Processor speed has approximately doubled every year since 1986. General purpose RISC processors are available with speeds in the range between 25 and 100 MIPS (million instructions per second). Graphics and image processing computers running at 300 MIPS (e.g. Vitec or Silicon Graphics) are available. These computers are programmable and can be used for a wide variety of tasks (Faust, et al., 1991a).

Concurrent with advances in the workstation or minicomputer processing power, advances are being realized in the PC class. The Intel 80486 and Pentium and the Motorola 68040 and PowerPC chip are providing the power to do tasks which have been traditionally reserved for mini or mainframe computers.

Almost all general purpose computers have been, until recently, single CPU computers. The availability of multi-processor machines is becoming more common. There are already several machines on the market such as the NCUBE/10, Sun Sparc10 and others.

These increases in power in off-the-shelf hardware mean that is no longer necessary to design and build single purpose image processing or geographic information systems. The days of the dedicated image processing station are rapidly disappearing.

2.1.2 Scanning Technology. Scanning is a widely accepted technique for the input data into a geoinformation processing system. It is often used to input map data from hardcopy to a GIS. Due to the unfortunate lack of operational digital aerial cameras, it is also the most common approach to acquiring data for a digital photogrammetric system. In order to achieve the resolution of an analog photo in a digital image, it must be scanned at approx. 10mm (2500 dpi) (Ehlers, 1991). While there are a number of scanners available which are capable of such resolutions (i.e. Zeiss Photoscan, HAI 100, Vexel), they are quite expensive.

For many applications, however, it is not necessary to have a digital product with the full resolution of an aerial photo. The field of desktop publishing has, to a large degree, driven most of the recent advances in desktop scanning. Users in this field have demanded, and received, higher resolution, color capability and lower cost. Manufacturers are now offering scanners that offer resolutions of 600 or 1200 dpi in format sizes suitable for scanning aerial photos which are as much as an order of magnitude cheaper than those offered specifically for photogrammetric use. For applications not requiring full photographic resolution, the results obtainable from desktop scanners offer suitable resolution and accuracy (Sarjakowski and Lammi, 1991).

2.1.3 Storage Capacity. The analysis of geographic data requires the storage of large amounts of data. Advances in disk drive and other mass storage technology have kept up with those in the processor field.

We have progressed from a point where 50 MB of hard disk storage required a cubic meter or more to being able to store 1.2 GB on a 3-1/2" hard drive that occupies about 0.2 cubic meters (in external configurations). The form factor for hard drives continues to decrease along with the increase in capacity. 3-1/2", 500 MB drives are common and smaller drives are being developed and introduced.

Other forms of magnetic media have also advanced. The 3-1/2", 1.2 or 1.4 MB floppy is the defacto standard. Magnetic tape systems such as the Exabyte are capable of storing 5 GB of data on a cartridge (with compression) that is roughly the size of an audio cassette.

Optical or magneto-optical (M-O) storage is further increasing storage capacity. CD-ROM and WORM (Write Once, Read Many) drives are commonly used for archiving data. M-O systems with read/write capabilities currently store 650 MB per disk and, with foreseen technical advances, this is expected to increase to 10 GB within the next 5 years. When coupled with juke box mechanism, users will have access to 100's of gigabytes of "near-line" data storage. Floptical disk systems are now providing read-write storage for 21 MB of data on 3-1/2" diskette.

2.1.4 Display Technology. In order for geographic information to be useful, it must be converted from its digital form to a visible product. The visible form that is most easily interpretable for the human user is that of a graphic product. Advances in display technology have brought the graphic display from a 256 X 256 pixel display capable of displaying pseudo-color images to a 1024 X 1024 pixel, "true color" and stereo display, with 2K X 2K or 4K X 4K displays becoming feasible.

2.1.5 Decreased Cost. What is interesting to note is that, with the exception of the extreme cutting edge of the advances mentioned above, this technology is commonly available in off-the-shelf systems. In addition, each advance in technology is accompanied by a corresponding decrease in cost. With each passing year, we can process more megabytes faster and cheaper.

2.2 Software Advances

2.2.1 Programming. In former years, most of the software for processing geographic data or images was written in FORTRAN or assembly language. Currently, this software is programmed in C or C++. This has meant, generally, an increase in both processing speed (in addition to hardware speed increases) and cross-platform portability. This is particularly important from the systems point of view since it allows programs developed, for example, on a workstation to be ported to the PC environment and vice versa.

This portability is also being enhanced by the general shift to the use of UNIX operating systems for program development and operation. While it was once considered impractical to run UNIX on a PC, this is no longer the case due to the advances in computing power. The recent introduction of Windows NT will bring further potential for cross-platform portability.

Efforts are also beginning to define a set of standards for geoinformation processing functions which can be implemented in a programming toolkit. This toolkit approach would be similar to those used by many of the windowing environments such as the X-11 toolkit (Faust, et al., 1991b).

2.2.2 Interface. Another important advance in geographic information processing has been the move towards the graphical user interface (GUI). Early systems required the use, first, of batch processing (essentially no user interface) and, then, a command line approach often requiring the memorization of cryptic commands and complex syntax. Today's software, for the most part, makes extensive use of a graphical user environment using windows, icons and buttons. In many cases, the interface is built using a set of standard interface tools such as those available in X-Windows, MS-Windows or the MacOS. This means that the 'look and feel' of the system is familiar to the user making learning and using the software easier.

Current trends indicate that, in the future, which interface a user prefers will have less impact on which software they use. Products and tools are being developed which will allow software developed for one interface to run under another. Examples include Suns WABI (for Windows software) and a recent agreement to port Apple's GUI to other platforms.

2.2.3 Integration. Where once there was a clear division between the various systems and software used for geographic information processing, we are seeing the boundaries becoming more and more blurred. In the past, dedicated image processing systems were used to extract information from satellite images and photogrammetric systems to do the same from aerial photos. These data might be transferred to a GIS for analysis or to a separate cartographic system for map output. This is no longer necessarily the case.

Even a cursory look at the literature will show that systems are being developed and marketed that combine the features originally available from a dedicated system. Image processing systems, such as those from Erdas or Intergraph, offer the capability to perform some geographic analysis and orthophoto production (formerly the realm of photogrammetry). On the other hand, geographic information systems, such as Arc/Info or ILWIS, offer image processing capabilities and hardcopy output, as do digital photogrammetric systems, such as those from Intergraph, Vexel, or Zeiss.

Many vendors are offering solutions to the "integration" of digital image processing and GIS. The definition of integration seems to be somewhat variable and the solutions range from the ability overlay vector GIS data on an image backdrop to the combination of "image processing, complete raster GIS modelling, and powerful topologic vector digitizing and editing" (Treadwell, 1991).

3. GEO[INFOR]MATICS - AN EVOLVING DISCIPLINE?

Recently, Advanced Imaging presented a survey of vendors that offer products for the acquisition, processing, display, storage and output of digital imagery. Its list contained about 1100 addresses, most of them from the United States (Advanced Imaging, 1993). While the majority of these specialized in areas such as desk top publishing or medical imaging, many of their products could be linked to photogrammetric, remote sensing or GIS applications. Surveys on GIS and remote sensing software that also demonstrate the integration of spatial data handling technologies were published by Parker (Parker, 1989) and Sader and Winne (Sader and Winne, 1991) and Ehlers (Ehlers, 1992).

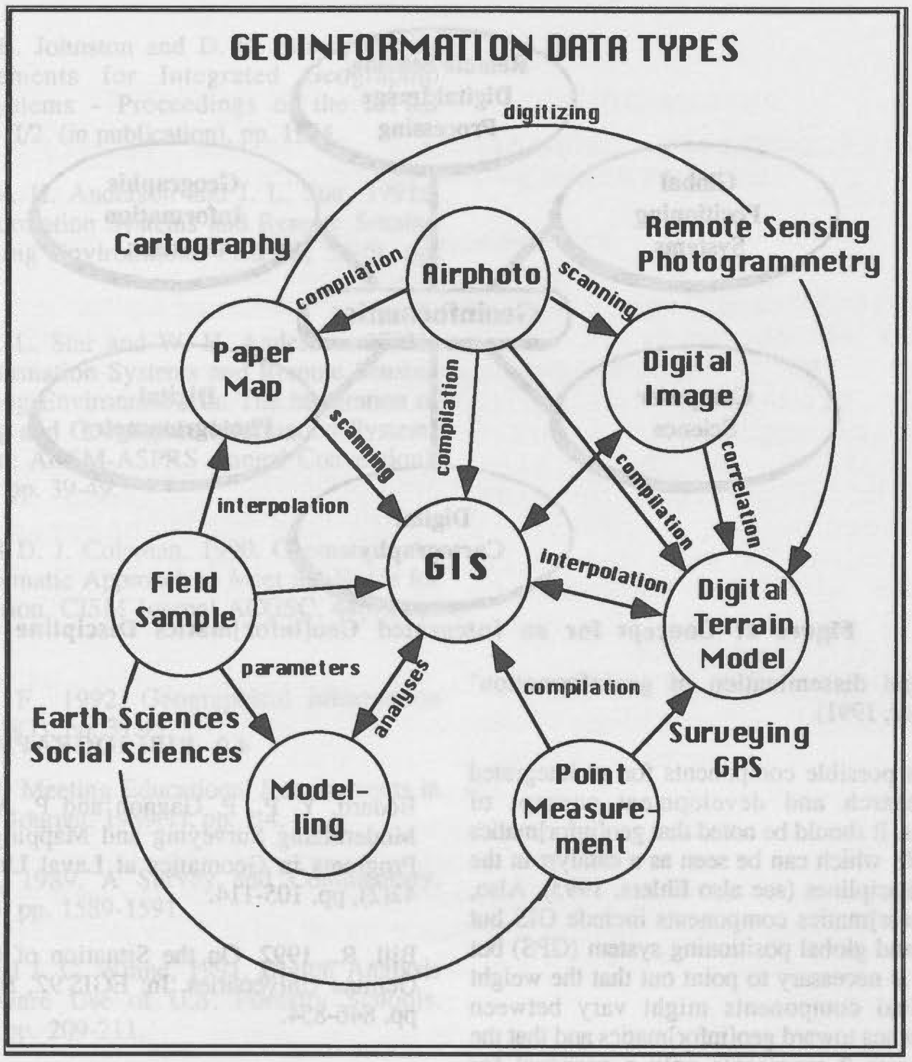


Figure 1: Geoinformation Data Types (arrows indicate possible conversion strategies)

Efforts to integrate spatial information disciplines have been primarily pushed by the advances of technologies. A number of proposals, however, have been made by scientists and scientific committees to define an integrated science or discipline rather than just looking into the progress of technology. For example, the surveying program of the University of Laval in Québec, Canada, developed a radically revised curriculum and degree program in 'Geomatics' (Bédard, et al., 1988) The term Geomatics was subsequently adopted by other Canadian organizations and institutions concerned with surveying and mapping (Groot, 1989).

Similar efforts to integrate specialized spatial data handling disciplines resulted in the formation of a Department of Geoinformatics at the International Institute for Aerospace Survey and Earth Sciences (ITC) in Enschede, The Netherlands (Ehlers and Amer, 1991, Groot, 1989). The Geoinformatics Department integrates photogrammetry, cartography, remote sensing, GIS/LIS and computer science (applied informatics). The ITC approach did not radically change an existing program in higher education but rather added a new and integrated geoinformatics component. Their approach has been primarily driven by the need to integrate multi-source geodata within a common system (Figure 1).

It has to be noted, however, that the field of geoinformatics is only an evolving one. Other approaches have been initiated from disciplines such as geography (Dollinger, 1992, Goodchild, 1992) or treat geoinformatics as a *subarea* of surveying (Bill, 1992).

In this process, geo[infor]matics has been defined as "...the science and technology dealing with the structure and character of spatial information, its capture, its classification and qualification, its storage, processing, portrayal and dissemination, including the infrastructure necessary to secure optimal use of this information" (Groot, 1989); as "...field of scientific and technical activities which, using a system approach, integrates all the means used to acquire and manage spatially referenced data as part of the process of producing and managing spatially based information" (Gagnon and Coleman, 1990); and as "...the art, science or technology dealing with the acquisition, storage, processing, production,

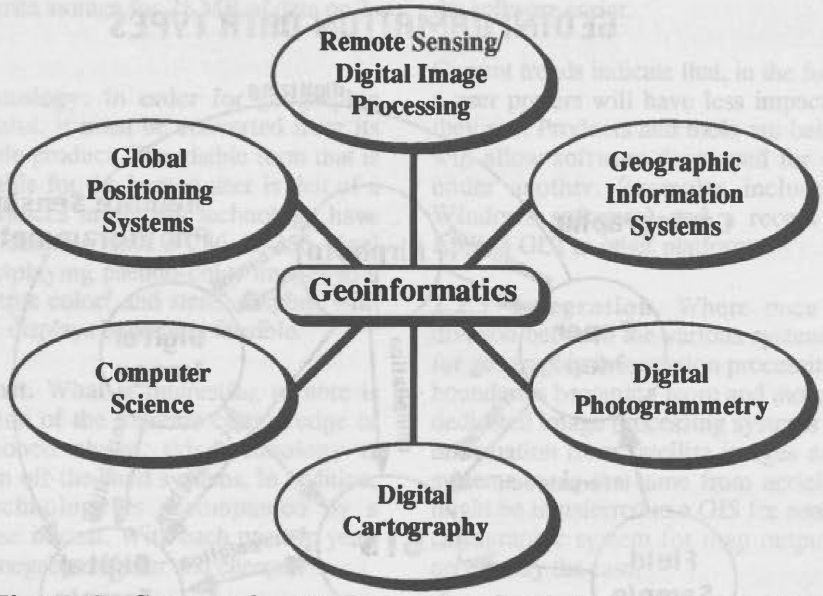


Figure 2: Concept for an Integrated Geo[infor]matics Discipline

presentation and dissemination of geoinformation" (Ehlers and Amer, 1991).

Figure 2 shows possible components for an integrated education, research and development concept of geo[infor]matics. It should be noted that geo[infor]matics is more than GIS which can be seen as a catalyst in the integration of disciplines (see also Ehlers, 1993). Also, here the geo[infor]matics components include GIS but not geography and global positioning system (GPS) but not geodesy. It is necessary to point out that the weight of the individual components might vary between different approaches toward geo[infor]matics and that the diagram of Figure 2 represents only a proposal for discussion.

4. CONCLUSIONS

In conclusion, we state that advances in computer hardware and software for GIS, image processing, digital photogrammetry and digital cartography are leading to the development of systems for integrated geoinformation processing. General purpose high resolution graphics workstations are being used for remote sensing image analysis, computer vision, digital photogrammetry, digital cartography and GIS. Boundaries between these fields will be less and less important, emphasis will be placed on integration aspects of these technologies. The emergence of standards and the ever-increasing processing, storage and display capabilities of modern RISC based CPUs is sees as a catalyst for the development of integrated approach to geoinformation handling.

4.0 BIBLIOGRAPHY

Bédard, Y. P., P. Gagnon and P. A. Gagnon, 1988. Modernizing Surveying and Mapping Education: The Programs in Geomatics at Laval University. CISMJ, 42(2), pp. 105-114.

Bill, R., 1992. On the Situation of GIS-Education at German Universities. In: EGIS'92, Munich, Germany, pp. 846-854.

Dollinger, F., 1992. Geoinformatik - Einige Randbemerkungen zur Entwicklung einer jungen Wissenschaft. In: Angewandte Geographische Informationstechnologie IV - Beiträge zum GIS-Symposium, F. Dollinger and J. Strobl (Eds.), Institut für Geographie der Universität Salzburg, Salzburg, Austria, pp. 7-10.

Ehlers, M., 1991. Digitization, Digital Editing and Storage of Photogrammetric Images. In: 43rd Photogrammetric Week, Universität Stuttgart, Stuttgart, pp. 187-193.

Ehlers, M., 1992. Hybrid Workstations in Geoinformatics: Requirements and Potential. EARSeL Journal Advances in Remote Sensing, pp.

Ehlers, M., 1993. Integration of GIS, Remote Sensing and Cartography: The Geoinformatics Approach. Geo-Informations-Systeme (GIS), 6(5), pp. 18-23.

Ehlers, M. and S. Amer, 1991. Geoinformatics: An Integrated Approach for Acquisition, Processing and Production of Geo-Data. In: EGIS '91, Brussels, Belgium, pp. 306-312.

Ehlers, M., J. B. Johnston and D. R. Steiner (Eds.), 1994. Requirements for Integrated Geographic Information Systems - Proceedings of the ISPRS Working Group II/2. (in publication), pp. 1994.

Faust, N. L., W. H. Anderson and J. L. Star, 1991a. Geographic Information Systems and Remote Sensing Future Computing Environment. PE&RS, 57(6), pp. 655-668.

Faust, N. L., J. L. Star and W. H. Anderson, 1991b. Geographic Information Systems and Remote Sensing Future Computing Environment. In: The Integration of Remote Sensing and Geographic Information Systems (Special Session: ACSM-ASPRS Annual Convention), Baltimore, MD, pp. 39-49.

Gagnon, P. and D. J. Coleman, 1990. Geomatics: An Integrated, Systematic Approach to Meet the Needs for Spatial Information. CISM Journal ACGSC, 44(4), pp. 377-382.

Goodchild, M. F., 1992. Geographical Information Science. IJGIS, 6(1), pp. 31.45.

Groot, R., 1989. Meeting Educational Requirements in Geomatics. ITC Journal, 1989(1), pp. 1-4.

Parker, H. D., 1989. A Survey and Commentary. PE&RS, 55(11), pp. 1589-1591.

Sader, S. A. and J. C. Winne, 1991. Digital Analysis Hardware/Software Use of U.S. Forestry Schools. PE&RS, 57(2), pp. 209-211.

Sarjakowski, T. and J. Lammi, 1991. Digital Photogrammetric Workstations for Urban GIS. In: Digital Photogrammetric Systems, H. Ebner, D. Fritsch and C. Heipke (Eds.), Wichmann-Verlag, Karlsruhe, pp. 274-288.

Treadwell, J. H., 1991. GIS News: GIS-Pak from Decision Images. PE&RS, 57(9), pp. 1131.

SYSTÈMES D'INFORMATION GÉOGRAPHIQUE INTÉGRÉS (SIGI) : ÉTAT DES SYSTÈMES ET RECHERCHES

Résumé

Cette communication présente l'état des progrès récents en matière de logiciels et de matériel destinés aux systèmes d'information géographique (SIG). Alors que les limites s'estompent entre les systèmes de traitement d'information géographique, d'analyse d'images, de photogrammétrie sur écran, de cartographie numérique, de visualisation et de visionique, les progrès se feront principalement en matière de mise au point de systèmes intégrés. Nous présenterons les exigences propres aux SIG intégrés (SIGI) surtout du point de vue des utilisateurs. Les efforts visant à instaurer une approche intégrée de la R et D et de l'éducation en «géo-informatique» pour la manipulation des données spatiales seront également présentés.

INTEGRATION OF REMOTE SENSING AND GIS.

Peter A. Woodsford

Laser-Scan Ltd,Science Park
Milton Road, Cambridge, CB4 4FY, U.K.
Telephone (44) (223) 420414, Fax (44) (223) 420044

Commission II - WG II/2

KEY WORDS: Raster/Vector Integration, Imagery, GIS, Knowledge-based Classification, Object-based Database, SAR, Landcover.

ABSTRACT:

Remote sensing has been operational since 1972 with the launch of Landsat-1. Since then, remotely sensed imagery has been used for a wide range of applications, from land use analysis to geological exploration. These applications have been driven by developments in image processing - such as classification and atmospheric correction techniques - and by the increasing spatial and spectral resolution of imagery, culminating in the many hundreds of bands of data available from the latest hyperspectral imaging spectrometry sensors.
The output from these processing systems has been primarily imagery - enhanced or classified, but usually in image format which is left for the specialist to interpret. This interpretation is often carried out within a GIS with the image data treated as a backdrop for feature extraction. This paper shows how the synergy of integrating vector and raster processing will increase the use of remote sensing in a range of environmental and other applications.

INTEGRATION DE LA TÉLÉDÉTECTION ET DU SIG.

RESUME:

La télédétection a été operationelle dès 1972 avec le lancement du Landsat-1. Depuis ce temp les images télédétectées ont etaient utilisées dans une gamme d'applications, dès l'analyse du terrain à l'exploration geologique. Ces applications ont été accionnées par des développements dans le traitement des images - telles quelles les techniques de classification et correction atmosphérique - et par l'increment de la resolution spatiale et spectrale des images, culminant dans des centaines des bandeaux de données accesibles du plus recent détection des images hyperspectrales du détecteur spectrometrique. Le materiel de sortie des systèmes de traitement a été essentiellement des images - rehausses ou classifies, mais d'habitude dans un format d'image qui laisse l'interpretation au specialist. C'ést interpretation très souvant exécuté dans le SIG avec l'image utilisé comme fond pour l'extraction de l'article. Cet papier desmontre comment la synergie d'integration du traitement du vector et raster augmentera l'usage de la télédétection dans une gamme des applications pour l'environnement et d'autres.

1. INTRODUCTION

The science of 'remote sensing' in its broadest sense has been developing since the 19th century with the invention of photography and the first aerial photographs taken from captive balloons. Throughout the 20th century, technological advances in a number of areas - the development of colour and infra-red sensitive films, aircraft and satellite platforms - enlarged the sphere of remote sensing with the development of applications such as mapping, geological exploration and meteorology making use of remotely sensed images. Remote sensing as it is currently practised, however, began with two major advances in technology - the launch of high resolution digital imaging systems (starting with Landsat-1 in 1972) and the development of minicomputers and image-display terminals in the 1970s.

With these advances, image processing systems rapidly evolved. By the early 1980s, a typical system would have functionality for image input, geometric correction, classification (supervised and unsupervised), image enhancement, convolution, arithmetic functions (eg. band rationing) and principal components analysis. These would be performed as batch or interactive operations, with special frame-store hardware used for image display.

The evolution continued throughout the 1980s, with an increased range of processing functions, data from new sensors (Landsat TM, SPOT, radar, airborne multispectral scanners), faster processors, higher resolution displays and user-friendly menu interfaces. Interfaces to vector data were provided by most systems, although with functionality largely limited to the overlay of vector data on imagery.

Geographic Information Systems (GIS), however, have developed from three largely separate origins. The Canada Geographic Information System (CGIS) was first proposed in 1963 and was designed for overlaying vector polygons of resource information for applications such as land use assessment. A different approach to the same problem was adopted by the Harvard Laboratory for Computer Graphics SYMAP system in the late 1960s, which performed overlay analysis with raster data. These systems were the precursors to the present-day systems designed primarily for data overlay.

A second origin to GIS was the development of vector-based digital mapping starting in the late 1960s. Digital mapping systems have evolved to the state where many map production flowlines now use computer technology for the manipulation of data. GIS systems have grown from digital mapping systems by the addition of a relational database for the storage of attribute data and by the provision of analytical functionality. The vector data model that has become dominant in this strand of GIS development, because line maps were familiar to the first generation of GIS users, involved in applications such as property management and the utilities. Vector mapping systems allowed them to replicate, in digital form, their conventional methods of working.

A third route for the development of GIS systems, also based on the vector data model, was by a similar evolution from the Computer-Aided Design (CAD) area.

GIS systems from each of these three origins have evolved towards a common basic range of capabilities - input, display and manipulation of both vector and raster data, raster DTM handling, topological structuring of vector data, vector and/or raster polygon modelling and analysis, and storage of attribute data in a relational database management system. Most systems today have the ability to perform basic image handling tasks, such as rectification and display as an image backdrop. Some systems allow an image processing package to share the same screen as the GIS, giving access to both image processing and GIS functionality. In addition, modern GIS system architectures can handle large raster and vector datasets in a seamless manner, so image processing can escape the constraints of image scenes and map data can escape the fetters of sheet boundaries.

2. CURRENT STATUS OF REMOTE SENSING/GIS INTEGRATION

It is useful to consider briefly why this strong division of vector and raster processing has occurred:

- The raster and vector data models are fundamentally different - one storing data as a regular array of values, the other storing data as a series of points, with implicit lines drawn between them (early vector data models were often simply a series of 'pen up', 'pen down' and 'move to XY' commands used for controlling a pen plotter). Even today, there is considerable uncertainty over how data held in these different models should be analysed together.
- Remote sensing and GIS analysis are also fundamentally different processes. With remote sensing, the source data is spectral information (i.e. reflected radiation in one or more spectral bands), and the main processing tasks are concerned with the labelling of each pixel such as 'wheat' or 'coniferous woodland'. With GIS, however, vector data usually represent the edges of objects, whose identity is known such as 'road edge' or 'woodland boundary'. Remote sensing is therefore primarily concerned with the production of data, whereas GIS is primarily concerned with the analysis of data for applications such as record management, route finding or locational analysis (although the analysis may well incorporate datasets derived by processing of remotely sensed data).
- Historically, remote sensing and GIS technologies developed as separate disciplines for both cultural and information technology (IT) reasons. Information technology, in particular display hardware, was strained to its limits to service the conflicting requirements of the two disciplines and so they grew up conditioned by bespoke environments, e.g. frame buffers for image processing. Computer hardware and software advances, stemming from RISC workstations, dramatic improvements in affordable memory and storage and Open Systems standards, have resulted in an IT environment capable of supporting both disciplines. The first generation of systems to exploit this have been in fact hybrids, with combined display of imagery and vector data, but discrete environments for the two sets of functionality.
- Due to the cultural separation of the two disciplines in the past, there has been a dearth of compatible data exchange formats.

However, a number of recent trends in remote sensing and GIS have been influential in bringing the two disciplines closer together:

- Environmental GIS applications are increasingly using data derived from remotely sensed images. These data may be an update or enhancement to a map (such as the mapping of forest boundaries) or the results of an image classification exercise.
- Remote sensing analysis is increasingly using GIS data. Examples of this include the overlay of vector data for presentation and location, and the use of DTMs for terrain correction of imagery and perspective visualisation.
- The rapidly improving price/performance of computers now enables both GIS and image processing to use the same workstation. For instance, image processing systems do not need to use specialised hardware or displays, and a standard workstation used for GIS has the performance required for image processing (both in power for processing large volumes of data and the graphics display capabilities for 24 bit colour images).
- The convergence of the two disciplines, the growing commonality of user requirements, and the beginnings of effective standardisation activities (*de facto* and *de jure*) are improving the situation on data compatibility and availability. In addition,

the impact of GPS techniques, and the ready availability of ground control points in an effectively standard reference system (WGS 84), greatly facilitates the registration, in practice, of remote sensing and GIS or map data.

A review of recent literature in the fields of GIS and remote sensing reveals that there is considerable interest in the integration of remote sensing and GIS. Sadler et al (1991), for instance, use census data in an attempt to improve the classification of urban areas. Census areas held in vector from are processed to generate a continuous surface raster dataset of population density value. This is then used as a pseudo 'image plane' in a classification process.

Janssen et al (1990) describe the use of a topographic map to classify an image on a 'per-object' basis. However, as the systems used (leading GIS and image processing packages) have no capabilities for integrated processing of this type, they rasterise map objects and then use these as a mask within a self-developed program for assessing the dominant class within the objects, from an image classified using a standard per-pixel classification process.

The conclusion inevitably reached, therefore, is that a level of integration substantially beyond co-existence of image processing and GIS is required. The statement:

> "The integration of image data into GIS is one of the great ideas whose time has come" (Faust et al, 1991)

is now extremely relevant, whilst the statement:

> "Remotely sensed images have been shown to be a cost effective means for update of GIS data" (Faust et al, 1991)

is clearly demonstrable. The chief operation that can be said to date to effectively integrate remote sensing and GIS is heads-up digitising. However, this has been due to the limited capabilities of software packages: further benefits of integration remain to be realised.

3. THE FUTURE FOR INTEGRATION OF REMOTE SENSING WITH GIS

The first important distinction to make is between remotely sensed data and raster data (which may, of course be, or be derived from, remotely sensed imagery). Most references to 'integration' refer to the integration of raster data with vector processing functionality.

Although integrated data processing can utilise data in both raster and vector forms, it is still necessary for the output data to be held in one form or another. For instance, a function to classify an image on a per-object basis will output data as an attribute of a vector object, and a function assessing land suitability for development might output data as a single raster 'suitability' value.

Processing functionality that genuinely integrates GIS data with remotely sensed data can be divided into two groups:

- Functions that enhance image processing operations.
- Functions that enhance vector operations. These generally make no distinction between the use of remotely sensed imagery and of any other raster dataset.

4. FUNCTIONS THAT ENHANCE IMAGE PROCESSING OPERATIONS.

There are three main areas where GIS data can be used to enhance standard image processing functions:

4.1 Geometric correction.

Conventionally, ground control point (GCP) co-ordinates are input to an image processing system as an image is being rectified. If the system can display and manipulate GIS data, however:

- GCPs can be extracted from a vector database (typically road junctions)
- Vector data can be overlaid on an image to verify the rectification process
- A DTM can be used to create an ortho-image by differential rectification

4.2 Image classification.

This is, perhaps, the main task performed by image processing system, and can be enhanced in a number of ways by the use of GIS data:

- The use of vector objects for training sets. Instead of identifying and digitising training sets on the image, an existing vector database (with attributes held in a relational database) can be used to select objects to be used for training the classifier.

- Object classification. If the image to be classified is comprised primarily of homogeneous objects (such as fields), it is more efficient to classify the image on an object-basis, with the objects defined by a vector dataset (ie. a single classification is given to each field, rather than to each pixel within the field, eg. Pedley and Curran, 1991). This gives the classification greater reliability (as untypical pixels - such as mixed pixels at the field edges - can be easily filtered out).

- Terrain illumination correction. Images with illumination differences caused by terrain can be corrected by applying an illumination model generated from a DTM (eg. Jones et al, 1988). (Incidently, if an image is to be used for an image-map, it may be necessary to use this technique to 'change' the illumination direction from south-east to north-west, to avoid the 'pseudo-relief' effect).

- Knowledge-based classification. The latest developments of classification algorithms (eg. Janssen and Middelkoop, 1992) begin to use GIS data for improving classification by applying roles. Typical GIS datasets that could improve the classification process include a DTM, rainfall, slope/aspect and existing land-use information.

4.3 Masking operations.

Vector objects can be used for a wide range of image processing functions to mask areas for processing. A simple example of this would be to use the coastline to mask out the sea or land when applying special processing algorithms to the other. Linear vector geometries representing boundaries of spatial changes in spectral response can be used to define fuzzy edges for mosaicing operations.

All of these functions have been investigated and shown to be worthwhile enhancements to current image analysis procedures. The main reason why these have not been more widely implemented, however, is the difficulty in building a comprehensive range of vector handling functionality into most image processing packages.

5. FUNCTIONS WHICH ENHANCE VECTOR OPERATIONS.

There are relatively few GIS functions that can be enhanced directly by the use of remotely sensed imagery. (However, data derived from imagery - such as a land-use classification - may be used routinely for a wide range of applications). GIS functions using remotely sensed imagery are:

5.1 Image-map backdrop.

Images are being increasingly used in areas of the world where adequate base mapping does not exist. This has the advantage of being cheaper than vector mapping, more up-to-date (and easily updated by acquiring the latest imagery), and often showing ground features that are not well represented by conventional map symbology. These maps are often quite adequate for locating 'foreground' information, such as pipes and cables.

5.2 Database update (eg. heads-up digitising).

This is the one area of image processing/GIS integration that is often cited. Imagery - particularly large scale aerial photography - is widely used for updating a wide range of GIS datasets from base maps to thematic overlays.

6. INTEGRATION OF GIS WITH RASTER PROCESSING

A number of other functions integrate vector data with raster data in general. These are important as they increase the general capabilities of GIS systems for handling raster datasets. The most common processing functions are:

6.1 Raster editing.

This is becoming increasingly important within a GIS, as raster base maps are used as a cheap and quick alternative to vector digitising the maps. Raster editing functions required include:

- Outlining areas to be edited by a vector polygon
- Creating new vector objects using sophisticated drawing tools (such as text font, positions, sizes, orientation), and then 'burning through' into the raster dataset.

These techniques are increasingly being considered by map publishing agencies either as an alternative to a fully vector database, or as the basis of a hybrid, evolutionary approach, spreading the costs of data capture in a manageable way.

6.2 Raster modelling.

This is the term used for the combination of two or more raster datasets using mathematical (+-*/) and boolean (AND, NOT, OR XOR) operators. More sophisticated packages include functions such as cost surfaces and zone operations (eg. find the perimeter of a zone of values), and have constructs such as DO and WHILE loops. These have the potential of integrating vector data within the same type of analysis constructs, such as 'IF (within polygon) AND (attribute = X) THEN...'. Although the output data structure will, in this case, still be a raster, vector data can be handled in the same way as a raster.

The ability to model using multiple rasters, where raster data can be from a variety of remotely sensed image bands, DTM and scanned map sources facilitates exploration of complex modelling scenarios such as estimation of cross country vehicle movement potential in fragile environments where limited conventional mapping may be available.

7. A STEP FORWARD

Real advances in the integration of remote sensing and GIS can only come with the development of integrated software environments where the display, interrogation, processing and analysis of raster and vector datasets is supported by a single system architecture which gives freedom to model and rapidly build applications according to the task in hand rather than being constrained by the functionality of a fixed function GIS or image processing system. The underlying database for such a system should be capable of integrated storage and indexing of any GIS or remotely sensed data, with a unified data access mechanism. Only with ready access to such heterogeneous data, modelling opportunities can increase in complexity.

Using the latest object oriented GIS technology Laser-Scan have developed such an integrated GIS and remote sensing system for the British National Space Centre (BNSC) under a contract placed and managed by the Defence Research Agency (DRA) Farnborough. Called IGIS, the system is built around a versioned object database which provides an intuitive way of holding spatial and non-spatial data. Both raster and vector continuous maps are supported. The object model provides a natural correspondence between real-world objects and the data that models them. As well as support for a variety of datatypes, IGIS also supplies references which enable objects to refer to each other. Collections of objects, with raster and vector components, can be combined into complex structures which mimic those found in the real world. Navigation of these structures provides a natural 'distributed' index enabling rapid access between related objects.

In addition, objects can exhibit behaviour ('methods'). Storing both the data and the mechanisms required to manipulate it together gives enormous flexibility for the development of complex environmental analysis and other complex data modelling application areas.

8. ENVIRONMENTAL APPLICATIONS AND AN INTEGRATED GIS

The integrated GIS described in the previous section is currently in use in three UK Department of Trade and Industry (DTI) funded pilot projects. These are currently running in National Environmental Research Council (NERC) institutes and are designed to test the capabilities of the initial system, to give Laser-Scan feedback on the suitability of the system for NERC research and to ensure that the IGIS will evolve to meet the wider needs of the environmental science community.

Each project focuses upon problems which require maximum integration of remotely sensed and more traditional GIS data types. The diversity of the projects indicates the flexibility required of an integrated GIS:

- BNSC (Remote Sensing Application Development Unit, RSADU) are investigating applications for the use of SAR data. Because of the speckle inherent in SAR images, radar backscatter parameters must be derived from a number of pixels rather than from one pixel location. This neighbourhood is usually of irregular shape outlining a relatively homogeneous area. This shape may be defined by a separate segmentation of the image or by a vector dataset digitised from published maps and updated by interactive editing. The data used comprises both a spatial dataset (the vector polygon boundaries) and non-spatial attribute data (for example, tree species, planting date and girth). The intimate connection between vector and raster datasets, such that there is a capability to update the spatial and non-spatial vector attributes directly from the remotely sensed imagery, is the key requirement of this project, with great potential for data collection for environmental management tasks, (Hinton and Baker, 1994).

- The Institute of Terrestrial Ecology (ITE), Monks Wood (Environmental Information Centre, EIC) are assessing the applicability of the IGIS for existing and planned GIS projects using the DTI funded Landcover map of Great Britain within ITE. Specifically, they are developing new analytical methods for future research projects such as the DoE Key Habitats project and the ITE Countryside Survey, both of which require land cover information as an integral part of their database. A particular area of interest is in the analysis of pattern in landscape from the raster and vector datasets of landcover. For example, the spatial distribution of woodlands as compared with soils, altitude and climate across the whole of Great Britain, and the distribution and the mean separation of woodlands in different size ranges across Great Britain. Work is concentrated on the development of pattern analyses within the IGIS that concentrate on the particular advantages of the system, employing the strengths of vector analysis for region definition and the ability of the system to combine raster and vector data.

- The third pilot project is being run by British Antarctic Survey, Cambridge, to exercise the capabilities of the IGIS to manipulate altimeter data from various sensors and to integrate data from other sources in order to generate a high quality Digital Elevation Model (DEM) of Alexander Island and Palmer Island in the Antarctic peninsula. The project will contribute to British Antarctic Survey's ability to produce accurate maps for environmentally sensitive areas where remote sensing provides the only available data.

9. SUMMARY AND CONCLUSIONS

After over 20 years of essentially separate development, remote sensing applications are beginning to recognise the advantages of integrating GIS datasets into image analysis. The integration of remote sensing and GIS has been the subject of one of the National Centre for Geographic Information and Analysis (NCGIA) research initiatives (Star et al, 1991). This has concentrated on topics such as data structures, parallel processing hardware environments and error analysis. Meanwhile, users attempting to integrate remote sensing with GIS have struggled on with inappropriate software, converting formats and 'integrating' data by displaying datasets on the same screen.

The real advancements in the integration of remote sensing and GIS can only come with the development of integrated software environments where the display, interrogation, processing and analysis of raster and vector datasets is part of the underlying system design. A part of this process will be the definition of the processing functions that are needed to integrate data. This development process is underway at Laser-Scan where an object oriented GIS has been developed which integrates the display and analysis of vector, raster, DTM, remotely sensed and non-spatial data.

REFERENCES:

Faust, N.L., Anderson, W.H. and Star, J.L., 1991. Geographic Information Systems and Remote Sensing Future Computing Environment. Photogrammetric Engineering and Remote Sensing, 57 (6):655-668

Hinton, J.C. and Baker, J.R., 1994. Image classification and analysis in an integrated GIS. Paper to be presented at EGIS/Mari Conference, Paris, April 1994.

Janssen, L.F., Jaarsma, M.N. and van der Linden, E.T.M., 1990. Integrating Topographic Data with Remote Sensing for Land-Cover Classification. Photogrammetric Engineering and Remote Sensing, 56(11):1503-1506

Janssen, L.F. and Middelkoop, H., 1992. Knowledge-based crop classification of a Landsat Thematic Mapper image. International Journal of Remote Sensing, 13(15):2827-2837.

Jones, A.R., Settle, J., J. and Wyatt, B. K., 1988. Use of digital terrain data in the interpretation of SPOT-1 HRV multispectral imagery. International Journal of Remote Sensing, 9(4):669-682.

Pedley, M.I. and Curran, P. J., 1991. Per-field classification: an example using SPOT imagery. International Journal of Remote Sensing, 12(11):2181-2192.

Sadler, G. J., Barnsley, M. J. and Barr, S. L., 1991. Information extraction from remotely sensed images for urban land analysis. Proceedings of EGIS '91, Brussels, Belgium, 2-5 April 1991, pp 955-964.

Star, J. L., Estes, J. E. and Davis, F., 1991. Improved Integration of Remote Sensing and Geographic Information Systems: A Background to NCGIA Initiative 12. Photogrammetric Engineering and Remote Sensing, 57(6):643-645.

FROM VIDEODISK GEOGRAPHIC INFORMATION SYSTEMS TO MULTIMEDIA EMPOWERMENT

Douglas Hadden

Marketing Director, Fifth Dimension Systems, 2500 Don Reid Drive, Ottawa, Ontario K1H 8P5

KEY WORDS: Geographic information system, Multimedia, Empowerment, Information centric

ABSTRACT

Despite developments in Geographic Information Systems (GIS) technology, potential benefits to organizations have not been fully realized. Only when GIS data and analysis have migrated from discrete systems and are available enterprise-wide, will end-users be fully empowered. Standardization in the multimedia world will be the conduit to bring GIS integration to large organizations. This trend began with the exploitation of multimedia and videodisk technology by GIS developers.

1. Introduction

In the mid 1980's, geographic information system developers began to exploit new technology in order to handle complex graphical information less expensively. Integrating optical disk technology and audio visual standards such as NTSC has become known as "multimedia".

One of the characteristics of both multimedia and GIS installations is that they tend to be specialized, discrete environments. The value of geographic information systems for rapid calculation and effective communication will only meet it's potential when it is integrated enterprise-wide on the desktop. Multimedia will be the vehicle to bring this about. This vision of GIS integration is not a new one, it has been articulated in varying ways by many people. The purpose of this paper is identify how multimedia is the vehicle to propel GIS corporate-wide, and that the technology is realizable today.

2. "Video Disk" GIS

Applications from simple colour maps with a minimum of data running on a personal computer to weather system simulations on super computers fit into the general category of GIS. The processing and memory requirements of higher-end GIS applications needed very powerful computing in the 1980's. The costs for integrating large amounts of raster and vector information made software expensive, reducing the penetration of GIS into organizations - even those that had a legitimate and immediate need for automation.

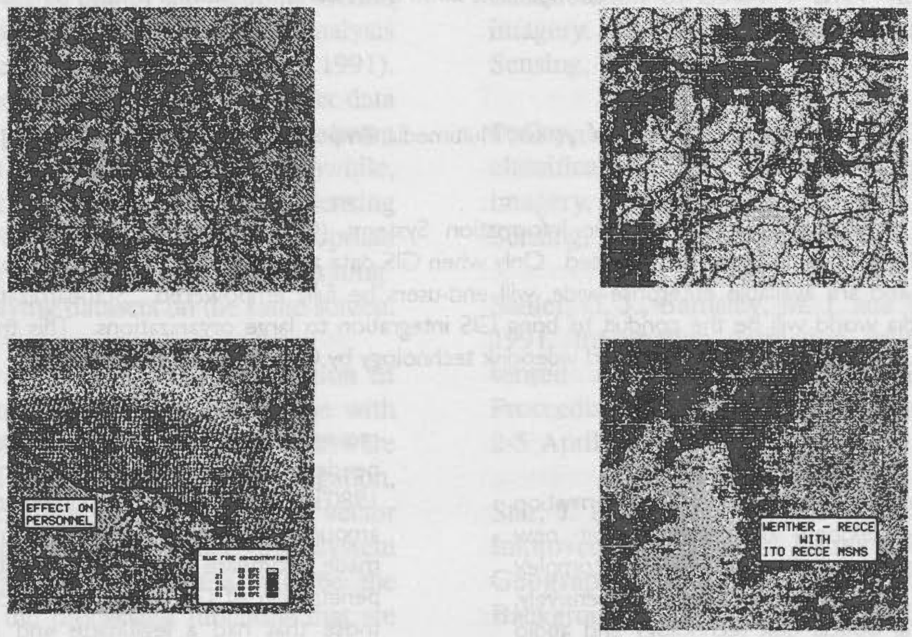
Since the 1960's when the value of digital processing was proven to be more effective than analogue processing, there has been a general tenet in the computer industry that systems are only elegant when all the information is digital and collected in a vector format with all sorts of attribute information. This has not always been a practical approach. In the case of GIS, developers have had to make decisions regarding the usefulness of information as it relates to processing and memory ramifications.

Video disk technology brought GIS to the personal computer without compromising the quality and amount of information that can be displayed. Although the information is digitally stored on a disk, it is in raster format and integrated with the computer through analogue inputs. A software database allows the linking of what are essentially pictures of hard copy maps with the computer. The system immediately shows the wealth of information available on hard copy maps to better communicate to the end-user.

At the outset of video disk GIS, map information was proprietary. There were no useful tools for taking mapping data from multiiple sources. The video disk allowed users to display a wider variety of information than was possible previously.

Microcomputers store additional vector and raster data. They can integrate with other databases such as the Digital Chart of the World to calculate distances and elevation in a transparent fashion.

Figure 1:



Video disk brings raster GIS to the personal computer.

The most effective application of the video disk GIS is command and control. In military and emergency preparedness requirements, full map information is crucial as is the ability to overlay vector information.

3. Raster and Vector

The advantage of the video disk is the quick display of complex raster information. This provided an immediate solution to the expense of processing this kind of data. However, raster provides very little assistance to the analysis of spatial data. Multimedia has been a solution for the presentation of vector data., as well. Compact disk technology has inexpensively brought shore line and elevation vector information to GIS applications. This data can be easily and inexpensively disseminated to a wide audience for both complex and entry-level requirements.

4. Multimedia

Multimedia applications range from a computer configured with a sound card to wide-area-networking video conferencing on demand. Multimedia applications process text, graphical, audio and video information together. The multiple source aspect tends to bring an immediate focus on hardware peripherals that have the ability to provide this data to the desktop with the appropriate performance and quality. In this aspect, the video disk met an immediate need to bring GIS data to the desktop computer. The software element of multimedia is that it is conceived to take advantage of multiple formats of data.

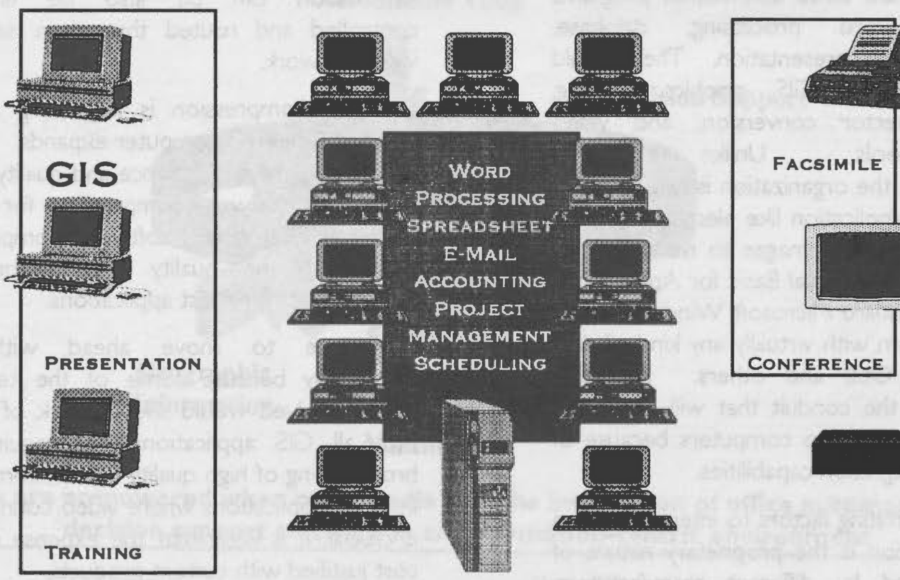
There are some key market perceptions that currently restricts organization's zeal to exploit the multimedia marketplace. One of the key barriers is that many people see multimedia as essentially a toy with little practical business use. Some see multimedia's potential but believe that it is not a mature technology and are waiting for new developments. Some see it as a new and complex environment that would be far more expensive to implement than traditional systems.

Although some key challenges remain, multimedia is practical, useful and can be implemented enterprise-wide. Multimedia development cycles are shorter and less expensive that traditional custom development.

The characteristics that are bringing multimedia in general, and GIS specifically, to the corporation are the adoption of standards, ingenious development by specialist companies and the exploitation of object-oriented programming.

In order to bring technology successfully to the marketplace, multimedia developers have created and adopted industry standards. Some of these standards relate to: video compression, video conferencing, CDROM, audio compression, colour output, user-interface and graphical data.

Figure 2



"Outside looking in"

Multimedia and GIS applications tend to be discrete, specialized and not integrated into the enterprise's information structure.

There are significant advantages in the adoption of standards other than the general benefit of protecting an investment. Many of these standards resolves bandwidth and processor-overhead problems associated with GIS graphical information. They allow for the use of "best of breed" software and hardware solutions. The most important aspect of the adoption of industry standards relates to the ability to integration data with a wide variety of applications.

Bill Gates of Microsoft has articulated a vision that he calls "information at your fingertips". The premise of information at your fingertips is that organizations will migrate from acquiring and using computer applications towards an information-centric environment. The characteristic of information centricity is that applications will be used as building blocks. Industry standards in multimedia are the link that will allow organizations to use GIS enterprise-wide.

Many organizations have reduced their budget for discrete GIS applications. The information centric model in which all sorts of integration is required to meet ultimate goals would appear to have an expensive price tag. As the public sector attempts to turn around deficits and the private sector attempts to be more competitive, management is hesitant to embark on new projects. Those organizations that have begun integrating GIS and other applications have found that multimedia information-centric development is far less expensive than originally anticipated.

With standards-based software and hardware, integrating a multimedia environment becomes an exercise in using proven off-the-shelf applications. These applications would include standard office automation programs such as word processing, database, spreadsheet and presentation. They would include specialized GIS, graphics handling, raster to vector conversion, and video interface tools. Unlike traditional development, the organization is not asked to re-invent an application like electronic mail in order to move GIS images to many people, Instead, tools like Visual Basic for Applications can link a standard Microsoft Windows-based E-Mail program with virtually any kind of data using DDE, OLE and others. Microsoft Windows is the conduit that will bring GIS data to most desktop computers because of its built-in integration capabilities.

One of the limiting factors to integrate GIS in the organization is the proprietary nature of data provided by different manufacturers. There have been tools developed to allow for the conversion and viewing of GIS data. The explosion of multimedia technology has fueled the need to use data from multiple sources. Therefore, one is more likely to find tools that read and write bitmap, CAD, and GIS, data in the multimedia world than in the GIS world. In this way, multimedia is driving the integration of GIS.

Icon-driven or object-oriented programming allows developers to see varying kinds of information as entities that can be invoked given conditions and rules. Authoring a multimedia environment involves the collection of data and the setting up of rules. Programmers are no longer faced with the task of command-line, non-portable low-level code. This allows an organization to prototype an idea quickly and gauge benefits in a fraction of the time of traditional computer programming.

5. Technical Challenges

There are a number of technical issues that make the implementation of a corporate multimedia environment difficult or expensive at this time. These issues relate to some missing pieces of the multimedia puzzle only. These include the difficulties relating to transmitting video and complex graphics through standard LAN wiring. This can be resolved today in a number of ways including hardware and software compression. Video information can be also be elegantly controlled and routed through a separate video network.

Software compression is improving as the power of desktop computer expands. At this time, there are performance and quality issues relating to software compression for video. However, still frame software compression will meet the quality and performance requirements for most applications.

Reluctance to move ahead with the technology because some of the technical issues involved would show a lack of vision. Not all GIS applications will require the broadcasting of high quality video information. In those applications where video connectivity is crucial, it is likely that the expense can be cost justified with current products.

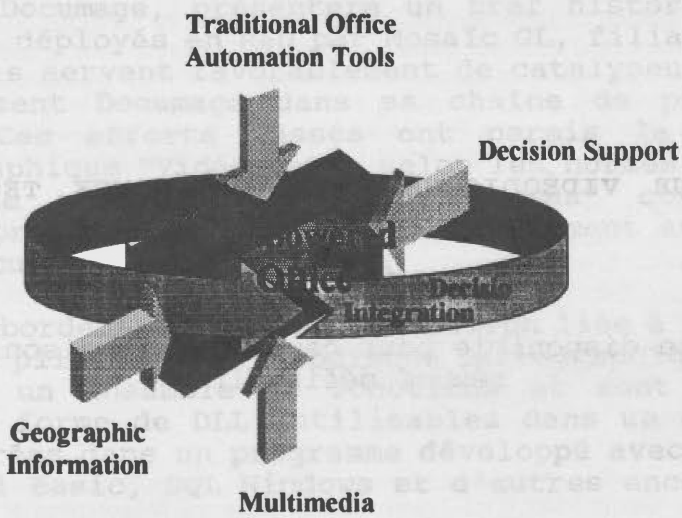
6. The Role of "Information"

GIS applications can be loosely categorized as those with a significant computational and analytical component and those with a more general corporate information component. A system charting weather conditions would be in the former category whereas a system keeping track of corporate facilities would be in the latter. This category will see significant long-term growth in GIS. That is not to say that there is no overlap. Base map and computational results are useful to others within the organization. In addition, the analysis of GIS data can be filtered down throughout the organization.

The "information age" with the development of the personal computer and multiple television channels saturates us with data. Everyone accepts that there is a relationship between access to information and improving decision making. However, the wealth of information tends to obscure decision making. That individual and important piece of information often does not jump off the printed page on a long report. Most organizations tend to restrict the information available to decision makers to the data that appears to be necessary. This is primarily summary information. Executive information software can also display this data in a graphical format to assist decision making.

The tactic of eliminating detail information for decision support is based on the premise that there is no appropriate method of massaging complex detailed data and making it useful. Multimedia communicates using all human senses except for smell. GIS techniques allows complex data to be displayed in easily-comprehended patterns. The key to empowering an organization is not to restrict the data, but rather to present that data in a useful and flexible way.

Figure 3:



Users are empowered when multimedia aids the integration of office automation and decision support software in an information-centric environment

7. Multimedia GIS "Office of the Future"

As a re-seller of computer aided design, geographic information, document management and multimedia systems, Fifth Dimension has encountered numerous installations which integrate GIS into a corporate system. There are many organizations on the cusp of implementing studies.

The need that is in the forefront of multimedia/GIS/corporate integration is operations centre integration. Computerized operations centres are used in both the public and private sectors for such needs as wide area network support and emergency planning. GIS had been a discrete tool used in these applications. It has only been recently that GIS has been integrated effectively for the operations centre and the organization as a whole.

A typical operations centre in the transportation sector would be functional twenty four hours of the day, seven days a week. Data relating to scheduling and weather conditions would be displayed to the information centre staff. That data would also be used by decision makers outside the operation centre when making budgeting decisions.

In the case of an emergency, for example, a train derailment of chemicals; the power of an integrated solution would be immediately visible. First, details about the derailment would be automatically E-Mailed or facsimiled to important individuals outside the operations centre. The president and public relations spokesperson would be immediately briefed of details prior to press interviews. Within the operations centre, the GIS integration would show a geographic display of the derailment. Based on weather conditions (from a GIS package) characteristics of the chemicals (from a chemical database), terrain (from a GIS package), population (from a GIS or database), and others; the organization can quickly react to reduce potential threats. Video information from television networks or previously recorded training tapes can be integrated.

8. Conclusions

Multimedia is the vehicle propelling GIS from the discrete application to the corporate desktop. GIS techniques for displaying complex information will be exploited by organizations to improve decision making.

ACKNOWLEDGEMENT

This paper benefited from the comments and assistance of Fifth Dimension colleagues: John Piraino, Dan McQuade and Ted Burnett.

DES SIG SUR VIDÉODISQUE À LA PUISSANTE TECHNOLOGIE MULTIMÉDIA

RÉSUMÉ

Traduction non disponible pour cause de livraison tardive du résumé définitif

VIDEODISK MAPPING IN CANADA

Eugenio Capodicasa

Chargé de projet de développement vidéodisques cartographiques
Documage, gestion et conversion inc.
3, Place Laval, Suite 090
Laval (Québec)
H7N 1A2
Tél. : (514) 629-6278; Téléc. : (514) 629-6735

RÉSUMÉ

Le titulaire de Documage, présentera un bref historique sur les efforts passés et déployés en R&D par Mosaïc GL, filiale de Gendron Lefebvre, lesquels servent favorablement de catalyseur aux efforts actuels que consent Documage dans sa chaîne de production de vidéodisques. Ces efforts passés ont permis la création du logiciel cartographique "Vidéoscan", selon les normes et standards de l'OTAN. Une présentation de "Vidéoscan" comportera une description des procédés de production entièrement automatisés et appliqués par Documage.

Le représentant abordera également la question liée à l'utilisation des librairies de pilotage du vidéodisque cartographique, celles-ci donnent accès à un ensemble de fonctions et sont présentement disponibles sous forme de DLL, utilisables dans un environnement Windows et intégrées dans un programme développé avec les langages comme C--, Visual Basic, SQL Windows et d'autres encore.

Ainsi l'auteur fournira, une information touchant le développement d'un nouveau support à la cartographie qu'est la technologie du CD-ROM, ce support à la publication et à la distribution de données géomatiques et topographiques, est prédestiné à s'imposer de plus en plus pour répondre à une demande dans le marché de la cartographie.

ABSTRACT

Mr. Eugino Capodicasa, will present a background on the R&D made by Mosaïc GL, Branch of Gendrn Lefebvre, that resulted in the creation of the technology "Vidéoscan" actually used by Documage to produce cartographic videodisk. This process complies with NATO standards. The presentation will describe the entirely automated process operations of used for the videodisk cartographic production.

We will also discuss the implementation of Libraries used in a Windows environment to pilot the videodisk. These libraries are available to developer of multi-media GIS as DLL.

Finally the emerging use of Digital Raster Graphic Product will be discussed in focus with the implementation of CD-ROM as a mean of storage of cartographic data.

SYSTÈME D'AIDE À LA PRODUCTION AUTOMATISÉE DE MOSAÏQUES D'IMAGES AÉRIENNES

André Leclerc, coordonnateur de projets
Photosur Géomat inc., 5160, boul. Décarie, bureau 300, Montréal, H3X 2H9, Canada

Valter Rodriguez, chercheur
Département de génie électrique, École Polytechnique, Université de Montréal
2900, boul. Édouard-Montpetit, B.P. 6079, succ. A, Montréal, H3C 3A7, Canada

Commission II, Groupe de travail II/III - Systèmes photogrammétriques numériques

MOTS CLÉS : Photomosaic, orthophotography, aerotriangulation, digital elevation model, digital photogrammetric system, digital image processing.

RÉSUMÉ

Un système permettant de produire de façon économique des mosaïques de photographies aériennes numériques est présentement en développement chez Photosur Géomat. Ce projet est réalisé dans le cadre de PROGERT (projet de recherche sur l'observation et la gestion des ressources terrestres), avec la participation de l'École Polytechnique de l'Université de Montréal. Le problème que solutionne le système est le suivant : Comment produire une mosaïque de photographies aériennes, géoréférencée, dans laquelle les photographies adjacentes coïncident le long des lignes de suture, sans avoir à produire un modèle numérique d'élévation du terrain au préalable, et en n'utilisant qu'un petit nombre de points d'appui.
Le système se compose d'une partie interactive et d'une partie de traitement en lot. La partie interactive permet un contrôle de la qualité au fur et à mesure que la mosaïque est construite; ce qui assure que le résultat final sera adéquat, une fois les images rééchantillonnées. Un prototype fonctionnel du système a été développé. Les résultats obtenus jusqu'à maintenant sont prometteurs et l'intégration du système dans un environnement de production devrait être complétée pendant l'été 1994.
Les produits obtenus constitueront une excellente alternative pour les applications de systèmes d'information à référence spatiale pour lesquelles les produits cartographiques standards sont trop coûteux.

ABSTRACT

A system for the economical production of digital aerial photo mosaics is presently being developed at Photosur Geomat. The project is undertaken as part of PROGERT ("projet de recherche sur l'observation et la gestion des ressources terrestres"), with the participation of the École Polytechnique of the University of Montreal. The system addresses the following problem : how to produce a georeferenced mosaic of aerial photos in which adjacent photos fit at their seam line, without previously having to produce a digital elevation model, and using only a limited number of control points.
The system has an interactive portion and a batch processing portion. The interactive portion allows for progressive quality control throughout the mosaic construction process, thereby ensuring adequate results after resampling. A functional prototype of the system has been developed. The results obtained so far are encouraging and the integration of the system in a production environment should be completed during the Summer of 1994.
The digital products obtained will constitute an excellent alternative for geographic information system applications for which standard cartographic products are too expensive.

1. INTRODUCTION

Plus les SIG (systèmes d'information géographiques) évoluent, plus on assiste à une intégration des données vectorielles (cartes) et matricielles (images). De même, plus la technologie du traitement numérique d'images devient abordable (scanners peu coûteux, diminution du coût de la mémoire sur disque, traitement de plus en plus rapide), plus le besoin pour des mosaïques numériques de photographies aériennes utilisées comme référence spatiale dans les SIG, croît.

Un tel exemple illustrant cette tendance est celui de l'utilisation d'un SIRS (système d'information à référence spatiale) pour la gestion d'un cadastre fiscal. La constitution d'un cadastre fiscal est souvent, dans les pays en voie de développement, une des premières étapes en matière de géomatique. Or, les produits cartographiques standards sont souvent trop coûteux. L'utilisation de mosaïques numériques de photographies aériennes géoréférencées pourrait donc être une alternative, dans la mesure où celles-ci pourraient être produites à des coûts relativement bas.

C'est ce besoin qui a amené Photosur Géomat à développer le système de production automatisée de mosaïques d'images aériennes. Ce système a été développé, dans le cadre de PROGERT (Projet de recherche sur l'observation et la gestion des ressources terrestres) (Poirier, 1991), en collaboration avec l'École Polytechnique de l'Université de Montréal.

Cet article présente d'abord le problème de la confection de mosaïques d'images aériennes à faible coût. Le fonctionnement du système de mosaïquage est ensuite décrit. Finalement, les résultats préliminaires des essais du prototype sont présentés.

2. PROBLÉMATIQUE

La première idée qui vient à l'esprit, lorsque l'on parle de mosaïquage de photographies aériennes, est l'utilisation de l'orthophotographie numérique, où l'on rectifie les photos numérisées en utilisant des MNÉ (modèles numériques de l'élévation du terrain). Cette technique a l'avantage d'être précise et, par conséquent, de fournir des mosaïques dans lesquelles les joints entre photos adjacentes coïncident bien. Le problème est que, pour les applications que nous visons ici, il est souvent trop coûteux de produire les MNÉ. En effet, la production d'un fichier d'orthophoto correspondant à la superficie couverte par une photo complète coûte entre 100 et 150 $. Ceci comprend la numérisation, la lecture et l'écriture des fichiers, la mesure des points d'appui et le rééchantillonnage. Par conte, la restitution photogrammétrique du relief et la production du MNÉ pour la superficie équivalente coûtent, ensemble, environ 7 fois plus cher, tout dépendant de la précision requise et du relief.

À l'opposé, si l'on n'utilise pas de MNÉ et que l'on procède par correction géométrique d'images standard, i.e. par méthode planimétrique, il est pratiquement impossible de faire coïncider les joints entre photos, à cause des déformations locales, induites par le relief. La figure 1 illustre ces déformations locales. Supposons que les points p' et p", respectivement situés sur les photos 1 et 2, correspondent à un même objet sur le terrain. Si l'on n'utilise pas de MNÉ, le point p', sur la photo 1 sera projeté sur le terrain (et dans la mosaïque) à la position P_0'; alors que sa position réelle devrait être P_v (position vraie). De la même manière, le point p", sur la photo 2 sera projeté à la position P_0". Tout dépendant du type d'objet physique auquel correspond le point P et de la distance P_0'-P_0" en pixel, le décalage pourrait être visible dans la mosaïque.

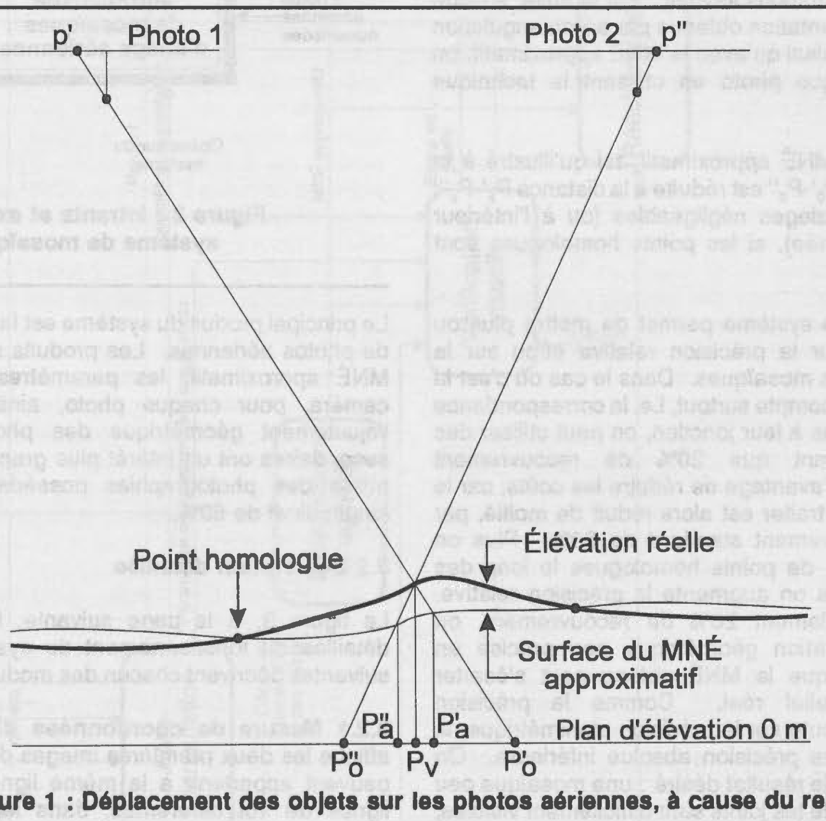


Figure 1 : Déplacement des objets sur les photos aériennes, à cause du relief

De là le besoin pour une méthode hybride, permettant de produire des mosaïques dans lesquelles les joints entre photos seraient difficilement visibles, sans toutefois posséder la précision des orthophotos, en tous points.

Les spécifications du système se résument donc par les éléments suivants :

- la confection de la mosaïque doit se faire à partir de quelques points de coordonnées cartographiques connues seulement;
- la partie du système où un opérateur doit intervenir ne doit pas contenir de traitement intensif, afin de réduire les coûts en temps personne;
- la correction des images par rééchantillonnage doit être effectuée en lot (sans intervention de l'opérateur);
- la partie interactive doit permettre une vérification des résultats suffisament fiable pour éviter une reprise éventuelle de la partie de traitement en lot;
- le système doit permettre de traiter au moins 5 x 5 photos à la fois;
- la précision et, par conséquent, le coût des mosaïques produites doivent être adaptables d'un projet à un autre et être optimisés en fonction des besoins du client.

L'analyse du problème nous a amenés à la conclusion que le modèle mathématique retenu devait tenir compte à la fois :

1) de paramètres géométriques globaux pour toute la mosaïque, servant à déterminer la position de chaque photo à l'intérieur de celle-ci (géoréférence);

2) de paramètres géométriques locaux, servant à modéliser les déformations locales des photos, près des lignes de suture entre photos.

3. DESCRIPTION DU SYSTÈME

3.1 Description générale

La méthode utilisée pour solutionner le problème est la suivante. Pour faire la liaison des photos entre elles, des points homologues sont mesurés à l'intérieur des zones de recouvrement. L'ajustement géométrique global des photographies est effectué au moyen d'une aérotriangulation par gerbe spatiale ("bundle adjustment") (Chapman, 1975). À partir des coordonnées de terrain obtenues par aérotriangulation pour chaque point homologue, on interpole un MNÉ approximatif. Celui-ci sert de modèle des déformations locales. Par la suite, à l'aide des paramètres d'orientation obtenus par aérotriangulation pour chaque photo, ainsi qu'avec le MNÉ approximatif, on rééchantillonne chaque photo en utilisant la technique d'orthophotographie.

En construisant un MNÉ approximatif, tel qu'illustré à la figure 1, la distance P_0'-P_0'' est réduite à la distance P_a'-P_a''; ce qui rend les décalages négligeables (ou à l'intérieur d'une tolérance donnée), si les points homologues sont bien choisis.

Selon les besoins, le système permet de mettre plus ou moins d'emphase sur la précision relative et/ou sur la précision absolue des mosaïques. Dans le cas où c'est la précision relative qui compte surtout, i.e. la correspondance des photos adjacentes à leur jonction, on peut utiliser des photos ne possédant que 20% de recouvrement longitudinal. Ceci a l'avantage de réduire les coûts; car le nombre de photos à traiter est alors réduit de moitié, par rapport à un recouvrement standard de 60%. Plus on augmente le nombre de points homologues le long des lignes de suture, plus on augmente la précision relative. Toutefois, avec seulement 20% de recouvrement, on obtient une configuration géométrique peu précise en altimétrie; de sorte que le MNÉ obtenu peut s'écarter passablement du relief réel. Comme la précision altimétrique se répercute sur la précision planimétrique, la mosaïque a alors une précision absolue inférieure. On obtient quand-même le résultat désiré : une mosaïque peu coûteuse dans laquelle les joints sont difficilement visibles, sans posséder la précision d'une orthophoto.

Pour augmenter la précision absolue, il suffit d'augmenter le recouvrement entre les photos. En utilisant des photos possédant un recouvrement standard de 60%, les points homologues peuvent être mieux répartis; ce qui stabilise l'aérotriangulation en altimétrie. De plus, ceci fournit une meilleure répartition des points pour l'interpolation du MNÉ. On obtient alors un MNÉ qui s'approche davantage du relief vrai; sans toutefois être comparable à un MNÉ obtenu par restitution photogrammétrique, bien entendu, à cause de la faible densité des points servant à l'interpolation.

Il demeure que le MNÉ approximatif a surtout pour fonction de modéliser les déformations locales entre photos. Dans un cas comme dans l'autre, la majorité des points homologues sont mesurés le long des lignes de suture entre les photos. La figure 2 montre une vue d'ensemble du système de production automatisée de mosaïques d'images aériennes. Les données d'entrée du système consistent principalement en des photograpies aériennes numérisées. Les données auxiliaires, servant au contrôle du système, sont les spécifications du projet et des coordonnées de terrain de quelques points d'appui. Ces coordonnées peuvent provenir de cartes de base existantes à petite échelle ou de mesures GPS. Un opérateur est nécessaire pour exécuter les tâches interactives et démarrer les traitements en lot.

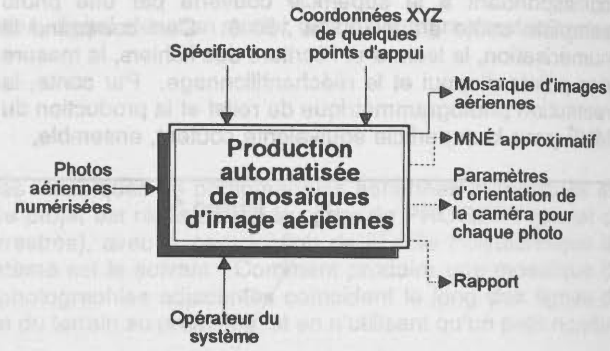


Figure 2 : Intrants et extrants du système de mosaïquage

Le principal produit du système est la mosaïque numérique de photos aériennes. Les produits secondaires sont : un MNÉ approximatif, les paramètres d'orientation de la caméra, pour chaque photo, ainsi qu'un rapport sur l'ajustement géométrique des photos. Ces produits secondaires ont un intérêt plus grand dans le cas où l'on utilise des photographies possédant un recouvrement longitudinal de 60%.

3.2 Description détaillée

La figure 3, à la page suivante, fournit une vue plus détaillée du fonctionnement du système. Les sections suivantes décrivent chacun des modules qui le composent.

3.2.1 Mesure de coordonnées d'image. L'opérateur affiche les deux premières images de la mosaïque. Elles peuvent appartenir à la même ligne de vol, ou à deux lignes de vol différentes; dans lequel cas, elles sont affichées l'une au-dessus de l'autre. Au moyen du curseur, il pointe d'abord les positions des marques fiduciales, puis celles d'une série de points homoloques. Si un point d'appui (de coordonnées de terrain connues) se trouve à l'intérieur des images affichées, celui-ci est identifié. Le nombre de points homologues à mesurer dépend de la précision désirée pour le joint entre les photos, ainsi que du relief en cet endroit.

Pour avoir une mesure immédiate de la précision obtenue, à la jonction des photos, l'opérateur appuie sur le bouton "calcul"; ce qui a pour effet d'exécuter l'aérotriangulation, l'interpolation du MNÉ, ainsi que de démarrer le module de vérification visuelle (section 3.2.4). En fonction du résultat obtenu, l'opérateur décide s'il ajoute ou non des points homologues pour densifier le MNÉ. Lorsque le résultat est satisfaisant, on passe à la photo suivante.

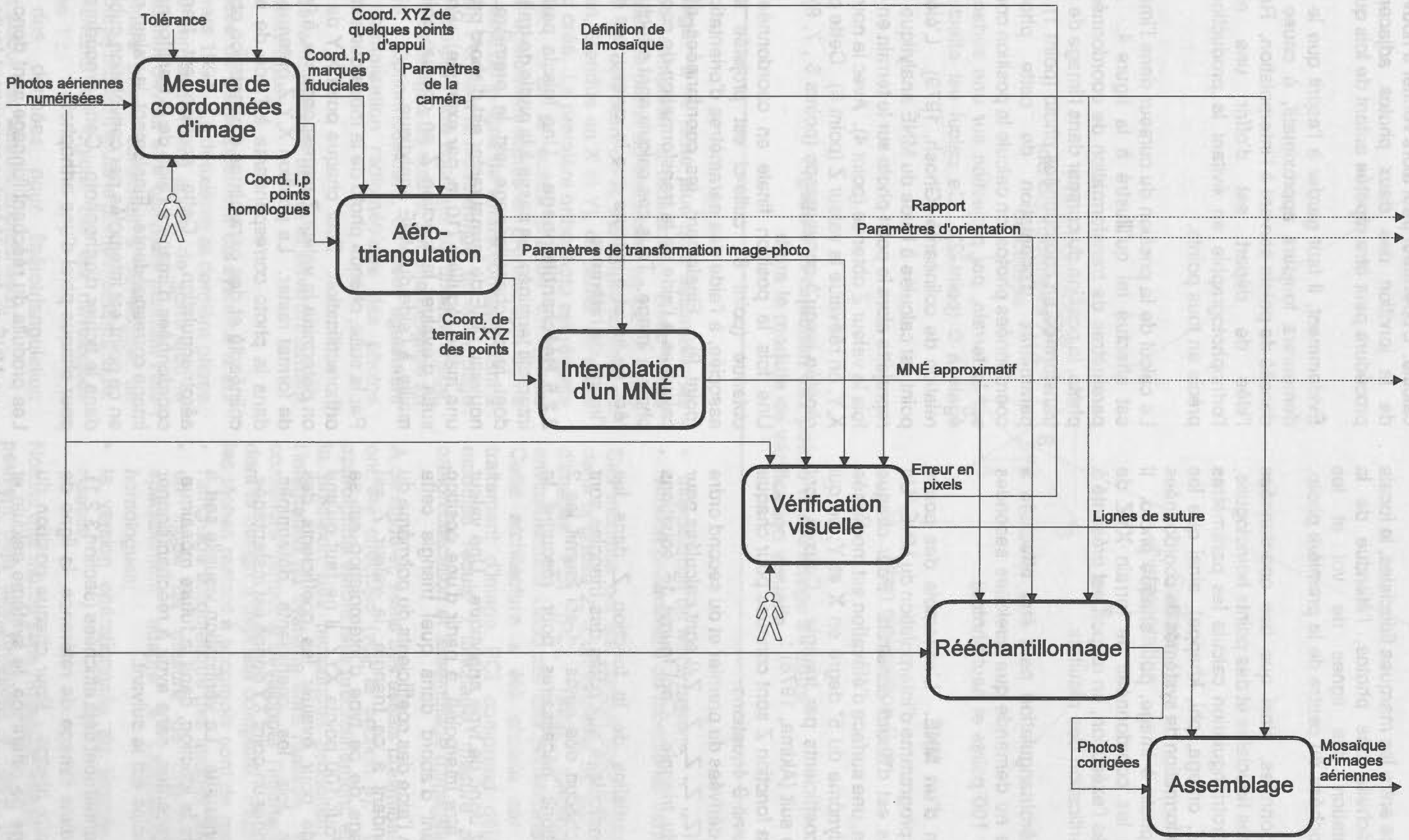


Figure 3 : Fonctionnement du système de mosaïquage

3.2.2 Aérotriangulation. Certains paramètres globaux du projet sont fournis par l'opérateur à l'entrée du système de mosaïquage et sont placés dans la base de données. Ces paramètres servent surtout à contrôler l'aérotriangulation.

Ce sont : la distance entre les marques fiduciales, la focale de la caméra, l'échelle des photos, l'étendue de la mosaïque, l'orientation des lignes de vol et les coordonnées approchées du centre de la première photo.

À l'aide de ces données, ainsi que des coordonnées d'image des marques fiduciales et des points homologues, le programme d'aérotriangulation calcule les paramètres d'orientation (X,Y,Z, omega, phi, khappa), ainsi que les paramètres de transformation de systèmes de coordonnées image à photo et photo à image, pour chaque photo. Il calcule également les coordonnées de terrain XYZ de chaque point. Après l'exécution, un rapport est présenté à l'opérateur pour vérification des résultats.

L'exécution de l'aérotriangulation peut être effectuée à volonté, puisqu'elle ne demande que quelques secondes (12 secondes pour 100 points et neuf photos).

3.2.3 Interpolation d'un MNÉ. L'ensemble des points XYZ est envoyé au programme d'interpolation du MNÉ. Un réseau de triangles est d'abord construit. Pour chaque facette des triangles, une surface d'élévation est modélisée au moyen d'un polynôme du 5° degré en X et Y; ce qui correspond à 21 coefficients par triangle. Ceux-ci sont déterminés comme suit (Akima, 1978) :

- Les valeurs de la fonction Z sont connues pour chaque vertex. Ceci fournit 3 équations.
- Les valeurs des dérivées du premier et du second ordre de la fonction Z (Z_x, Z_y, Z_{xx}, Z_{xy}, Z_{yy}) sont calculées pour chaque vertex du triangle. Ceci fournit 15 équations additionnelles.
- Les dérivées partielles de la fonction Z dans les directions perpendiculaire aux côtés des triangles sont calculées pour chacun des côtés. Ceci fournit les 3 dernières équations nécessaires pour résoudre le système.

Le MNÉ obtenu est donc de type analytique. Une valeur d'élévation Z peut être interpolée, à partir d'une position X,Y, en déterminant d'abord dans quel triangle cette position se situe. À l'aide des coefficients du polynôme du 5° degré correspondant à ce triangle, la valeur Z est calculée. L'avantage de ce type d'interpolation est sa grande rapidité. Pour 100 points XYZ, il ne faut qu'une fraction de seconde pour évaluer les coefficients des polynômes. Une fois les coefficients déterminés, l'interpolation du Z pour un point XY donné est instantanée.

3.2.4 Vérification visuelle. La vérification visuelle sert à évaluer la qualité de la jonction, dans la future mosaïque, de deux photos adjacentes, sans avoir à rééchantillonner les photos. Le principe est le suivant.

Deux images adjacentes sont déjà affichées (section 3.2.1). L'opérateur trace, dans l'image de référence, la ligne de suture entre les deux photos. Pour chaque position du curseur dans l'image de référence, le système calcule et affiche la position correspondante dans l'image associée. Si les positions correspondent - par exemple, le long d'une route - le joint sera parfait dans la mosaïque. Si elles ne correspondent pas, l'écart en pixels visible sera le même dans la mosaïque. L'opérateur peut alors décider d'ajouter des points homologues aux endroits où la tolérence est dépassée. En réexécutant les calculs, un nouvel ajustement, comprenant les nouveaux points homologues, est effectué par aérotriangulation et un nouveau MNÉ calculé. L'opérateur peut alors vérifier à nouveau la qualité de la jonction des deux photos adjacentes. Cette procédure peut être répétée autant de fois que l'on désire.

Évidemment, il faut garder à l'esprit que le MNÉ obtenu demeurera toujours approximatif, à cause de la faible densité de points servant à l'interpolation. Rappelons que l'idée de départ est d'offrir une alternative à l'orthophotographie en évitant la production d'un MNÉ précis en tous points.

Le calcul de la position du curseur dans l'image associée est effectué tel qu'illustré à la figure 4. À l'aide des paramètres de transformation de coordonnées d'image à photo, la position du curseur dans l'image de référence est transformée en coordonnées photo (point 1). À l'aide des paramètres d'orientation de cette photo, et des coordonnées photos, on calcule la position correspondante sur le terrain, par projection sur une surface d'élévation égale à 0 (point 2). Ce calcul est effectué suivant la relation de colinéarité (Ghosh, 1979). L'élévation en ce point est calculée à l'aide du MNÉ analytique (point 3). On reprojette alors le point photo sur le terrain en utilisant cette fois la valeur Z obtenue (point 4). Avec la nouvelle position X,Y, on réévalue la valeur Z (point 5). Cette procédure est répétée jusqu'à convergence (points 6 , 7, 8).

Une fois la position finale en coordonnées de terrain obtenue (point 8), celle-ci est projetée sur la photo associée, à l'aide des paramètres d'orientation de celle-ci (point 9). Finalement, les coordonnées d'images sont calculées à l'aide de la transformation de coordonnées de photo à image. Tous ces calculs sont effectués en temps réel, au fur et à mesure que le curseur est déplacé dans l'image de référence.

3.2.5 Rééchantillonnage. Une fois la partie de travail interactif terminée, on passe à la partie de traitement en lot, dont le rééchantillonnage est la première étape. Un nouveau MNÉ de format raster est d'abord produit, suivant une maille régulière (10 m, par exemple). Pour ce faire, il suffit d'évaluer la fonction Z en chaque point X,Y de la maille, à l'aide du MNÉ analytique.

Par la suite, chaque photo de la mosaïque est corrigée par orthorectification. Pour chaque point X,Y de l'orthophoto, on détermine la valeur Z correspondante, à l'aide du MNÉ de format raster. La position X,Y,Z obtenue est projetée dans la photo correspondante à l'aide de la relation de colinéarité et des paramètres d'orientation obtenus par aérotriangulation. Cette position est transformée en coordonnées d'image à l'aide de la transformation photo à image correspondante. Finalement, la valeur radiométrique en ce point est interpolée par convolution cubique et écrite dans le fichier d'orthophoto. Cette procédure est répétée pour chaque pixel d'une orthophoto.

Les produits du rééchantillonnage sont donc :

- un MNÉ de format raster recouvrant toute la mosaïque (celui-ci n'est pas indiqué sur la figure 3, pour ne pas l'alourdir);
- une série d'orthophotos prêtes à être assemblées dans la mosaïque.

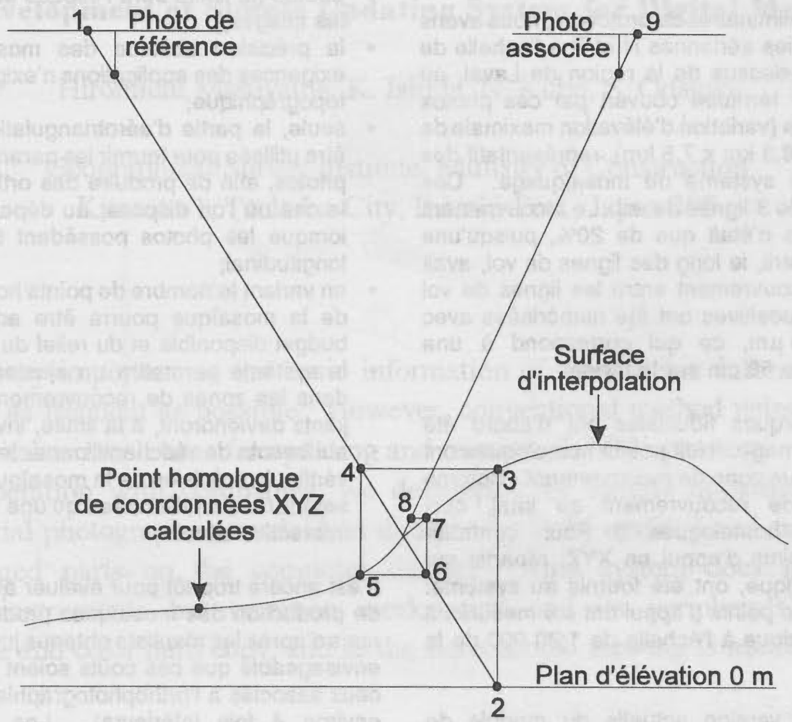


Figure 4 : Projection des points d'une image à l'autre dans le module de vérification visuelle

3.2.6 Assemblage. Il ne reste plus qu'à assembler les orthophotos pour compléter la mosaïque. Cette étape constitue la deuxième partie du traitement en lot. À l'aide des paramètres de définition de la mosaïque fournis au début (résolution, étendue en X et Y), un fichier de sortie vide est d'abord créé. La première orthophoto est réécrite dans ce fichier. Par la suite, chaque orthophoto additionnelle est également copiée en prenant soin de ne pas réécrire au-delà de la ligne de suture correspondant à cette orthophoto. Pendant le mosaïquage, un estompage des niveaux de gris peut être appliqué sur chaque image, de part et d'autre de la ligne de suture, sur une distance en pixels prédéfinie. Optionnellement, on peut également effectuer une normalisation radiométrique des photos adjacentes.

3.3 Stade de développement actuel et environnement de production futur.

La première phase de développement du système a été complétée en mars 1994. La deuxième et dernière phase a été démarrée immédiatement après; de sorte que le système opérationnel devrait être complété pendant l'été 1994. La première phase consistait à développer un prototype fonctionnel incluant toute la partie mathématique. Les modules suivants de logiciel ont été développés pendant la phase 1 :

- préparation des données pour l'aérotriangulation (formatage et calcul des valeurs approchées);
- aérotriangulation;
- toutes les transformations de coordonnées (image à photo, photo à terrain, terrain à photo);
- interpolation du MNÉ analytique;
- interpolation du MNÉ raster;
- version préliminaire du module de vérification visuelle.

Ces modules ont été conçus pour pouvoir être intégrés, dans la phase 2, à un système de traitement numérique d'images PCI (PCI, 1993), fonctionnant dans un environnement Unix, sur une plateforme IBM Risc/6000. Cette architecture a été choisie car le système de traitement d'image PCI comprend déjà plusieurs des modules requis : affichage, mesure de coordonnées, orthorectification et assemblage d'images.

À ce stade-ci, seuls des essais préliminaires du prototype ont été effectués (voir la section 4). Des essais plus approfondis seront faits prochainement, afin de déterminer la précision des mosaïques produites. Par la suite, l'intégration du système sera réalisée. Cette intégration consistera surtout à développer l'interface avec l'usager, dont le design a déjà été complété. Plus spécifiquement, les tâches restant à accomplir sont de mettre au point :

- l'affichage automatique des images par paires;
- un mode de mesure de coordonnées d'images spécialement adapté à l'acquisition des points homologues;
- la version opérationnelle du module de vérification visuelle.

Notons que même sans l'interface usager, des mosaïques peuvent déjà être produites à l'aide des résultats du prototype.

4. ESSAI DU PROTOTYPE

Pour faire les essais préliminaires du prototype, nous avons utilisé neuf photographies aériennes N et B à l'échelle de 1:10 000, acquises au-dessus de la région de Laval, au nord de Montréal. Le territoire couvert par ces photos présentait un relief faible (variation d'élévation maximale de 53 m sur une zone de 8,3 km x 7,6 km), représentatif des territoires visés par le système de mosaïquage. Ces photos faisaient partie de 3 lignes de vol. Le recouvrement longitudinal des photos n'était que de 20%, puisqu'une photo sur deux seulement, le long des lignes de vol, avait été numérisée. Le recouvrement entre les lignes de vol était de 30%. Les diapositives ont été numérisées avec une résolution de 50 µm, ce qui correspond à une dimension des pixels de 50 cm sur le terrain.

Les positions des marques fiduciales ont d'abord été mesurées sur chaque image. Huit points homologues ont été mesurés dans chaque zone de recouvrement. Comme il y avait 12 zones de recouvrement au total, ceci représente 96 points homologues. Pour contrôler l'aérotriangulation, 5 points d'appui en XYZ, répartis sur l'ensemble de la mosaïque, ont été fournis au système. Les coordonnées de ces points d'appui ont été mesurés à partir de la carte numérique à l'échelle de 1:20 000 de la région.

Nous avons utilisé la version actuelle du module de vérification pour vérifier la correspondance des photos dans les zones de recouvrement. Près des endroits où des points homologues avaient été mesurés, les positions indiquées par le curseur dans les photos adjacentes correspondaient, dans tous les cas. Ceci démontre que les algorithmes développés fonctionnent comme prévu.

Entre les points homologues mesurés, des écarts pouvant aller jusqu'à 5 pixels ont été observés, dans certains cas. Ceci est tout à fait normal et est dû au fait que le relief change à ces endroits. C'est d'ailleurs la fonction du module de vérification d'aider l'opérateur à identifier les endroits où des points homologues doivent être ajoutés. Lorsque l'intégration sera complétée, il sera facile d'ajouter les points au fur et à mesure.

Des essais plus approfondis seront effectués sous peu, pour mesurer la précision des mosaïques produites. Différentes mosaïques seront produites en faisant varier le nombre et la répartition des points d'appui et homologues. Des photos possédant un recouvrement longitudinal de 60% seront utilisées, en plus de photos avec 20% de recouvrement. Les mosaïques seront comparées à une mosaïque d'orthophotos produite avec des données altimétriques provenant d'une stéréorestitution. Par la même occasion, on évaluera le rendement du système dans des cas pratiques.

5. CONCLUSION

Le prototype fonctionnel d'un système d'aide à la production automatisée de mosaïques d'images aériennes a été développé et testé.

Le système opérationnel, qui sera complété au courant de l'été 1994, aura les particularités suivantes :

- nul besoin de produire un MNÉ au préalable, comme avec l'orthophotographie;
- il n'y aura théoriquement pas de limite sur le nombre de photos (la limite sera l'espace sur disque pour contenir les images);
- la précision absolue des mosaïques répondra aux exigences des applications n'exigeant pas une précision topographique;
- seule, la partie d'aérotriangulation du système pourra être utilisée pour fournir les paramètres d'orientation des photos, afin de produire des orthophotographies, dans le cas où l'on dispose, au départ, d'un MNÉ précis et lorsque les photos possèdent 60% de recouvrement longitudinal;
- en variant le nombre de points homologues, la précision de la mosaïque pourra être adaptée en fonction du budget disponible et du relief du terrain;
- le système permettra un ajustement local des photos dans les zones de recouvrement, de tel sorte que les joints deviendront, à la limite, invisibles;
- nul besoin de rééchantillonner les images pour pouvoir vérifier la précision de la mosaïque; puisque celles-ci ne seront rééchantillonnées qu'une fois la partie de travail interactive terminée.

Il est encore trop tôt pour évaluer avec précision les coûts de production des mosaïques produites avec le système, mais d'après les résultats obtenus jusqu'à maintenant, il est envisageable que ces coûts soient nettement inférieurs à ceux associés à l'orthophotographie numérique (peut-être environ 4 fois inférieurs). Les mosaïques devraient répondre aux besoins des utilisateurs de SIRS qui recherchent des données peu coûteuses servant de référence spatiale.

REMERCIEMENTS

Les auteurs tiennent à remercier Madame Joanne Chaput et Monsieur Jean-François Meunier, tous deux de Photosur Géomat, pour leur contribution à ce projet.

RÉFÉRENCES BIBLIOGRAPHIQUES

Akima, Hiroshi; 1978. A Method of Bivariate Interpolation and Smooth Surface Fitting for Irregularly Distributed Data Points. ACM Transaction on Mathematical Software, Vol. 4, No. 2, June 1978, Pages 148-159.

Chapman, Michael A.; 1979. A Procedure for a Realistic Calibration of a Specific Close Range Photogrammetric System. Master's degree thesis, Department of Geodetic Science, Ohio State University.

Ghosh, Sanjib K.; 1979. Analytical Photogrammetry. Pergamon of Canada Ltd., 150 Consumers Road, Willowdale, Ontario, M2J 1P9, Canada.

PCI Inc., 1993. Using PCI Software, volume I and II. Version 5.2 EASI/PACE. PCI Inc., 50 West Wilmot Street, Richmond Hill, Ontario, Canada, L4B 1M5.

Poirier, Sylvain; 1991. Projet de recherche sur l'observation et la gestion des ressources terrestres (PROGERT). Télédétection et gestion des ressources. Comptes rendus du VII[e] congrès de L'Association québécoise de télédétection, pages 31-35.

Development of Simple Updating System for Digital Maps

Hiromichi Maruyama, K. Ishida, N. Kubo, S. Odagiri

Geographical Survey Institute, Ministry of Construction
Kitasato-1, Tsukuba-City, Ibaraki-Pref., Japan 305

ABSTRACT

Digital maps are very important as the basic information of Geographic Information System, and should be updated as frequent as possible. However, conventional method using stereo plotter and stereo photographs takes much time for updating and requires skillful operators.
Therefore, in cooperation with companies, we developed the simple updating system in which scanned digital aerial photograph is overlaid on digital map to be updated and an operator finds and updates the changed parts on the computer display. This system does not require special photogrammetric instruments but common workstation and its peripherals, such as scanner, cartridge disk drive and color hard copy, and as the stereoscopic viewing is not required, anyone can handle it easily.

KEY WORDS: Map revision, Digital aerial photograph, Digital map

1. INTRODUCTION

Maps should be revised according to the changes of earth surface such as the construction of new roads, buildings or the development of golf links. So far, when map revision is conducted, geometric accuracy has been mainly stressed on. Therefore three dimensional measurement using stereoplotter, which requires the special instruments, skillful operator, is applied for map revision.
Recently, the digital maps have come to be often generated and used in the GIS environment. This means, in order to increase the reliability of the results of analysis of the GIS operations such as the extractions of some map features or the overlay with other layers, the contents of map should be coincide with real situation more than before.
Based on these backgrounds, simple updating system for digital maps, which does not require special instruments and enable anyone to revise maps if digitized aerial photograph are available, has been developed by the cooperative work between Geographical Survey Institute (GSI) and companies. In the following, the basic design of updating system is summarized and the system designed for 1:25,000 scale map, which has been developed by GSI, and the system for 1:2,500 scale map, which has developed by companies, are described.

2. SYSTEM OVERVIEW

Simple updating system should be the system by which anyone can revise digital map easily. In order to attain this concept, the measurement by stereoscopic view should not be used and the extraction of changed parts should be able to be done easily. Therefore, in this system maps are revised manually using mouse after single digital photograph is overlaid on the digital map to be revised and the changed parts are recognized visually.
Figure 1 shows typical hardware configuration of simple updating system.

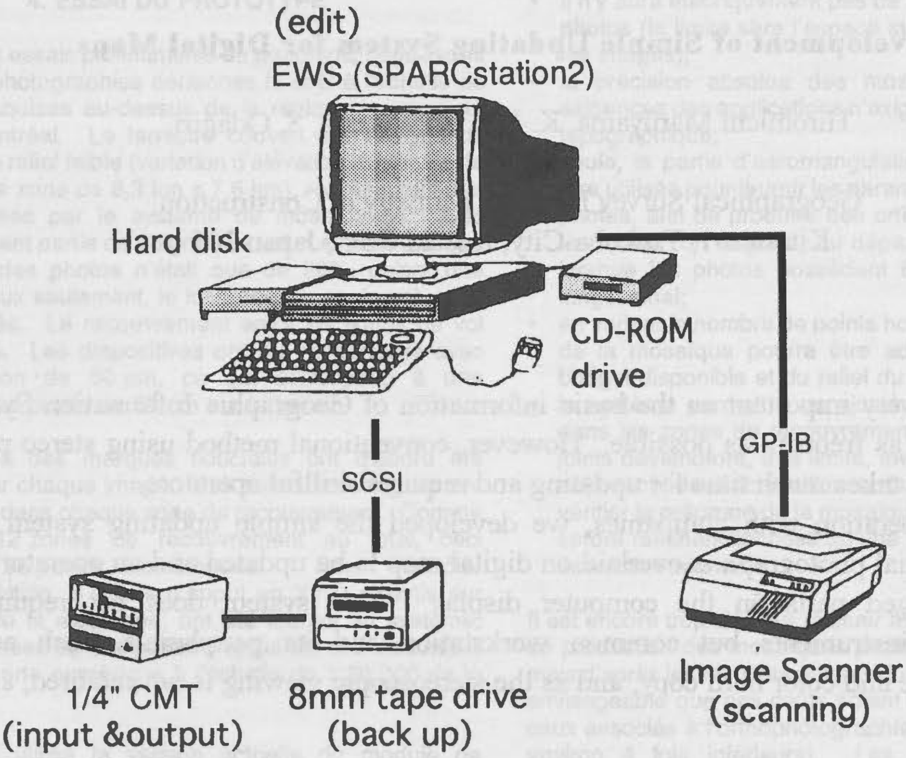


Figure 1 : Typical hardware configuration of simple up-dating system

-scanner used for digitizing aerial photographs,
-1/4" cartridge disk drive used for loading and storing the digital map,
-workstation used for overlaying, extracting changed parts and editing,
-8 mm tape drive for back up

Using these hardware, updating work is done in the following flow.
a. digitization of aerial photograph
b. overlaying aerial photograph on map
c. visual extraction of changed parts
d. edit

Computer programs corresponding to these works a,b,c,d are required. Among these, the program for overlaying is essential in this system.

3. FLOW OF UPDATING WORK

3.1 Digitization of aerial photograph

As quality of digitization of aerial photograph influences visual interpretation very much, scanning should be properly conducted according to the quality of final result required. There are two types of scanner, transparent type and reflective type. Transparent type, which can attain high geometric accuracy but is expensive, is for film, and reflective type, which is common in small size scanner, is for print. Therefore, reflective type scanner is better from the point of simple updating system in case high geometric accuracy is not required.

In case geometric accuracy is important, aerial photograph should be digitized in finer than 25 μm by transparent type scanner. However, if scanning pitch is too fine, data size of scanned photograph become too huge.

3. 2 Overlaying aerial photograph on map

There are three types of methods for overlaying used in this system. Those are as follows.

a. method using orthophoto

b. method using digital map transformed into central projection

c. method using affine transformed photograph

Principle of these methods are shown in Figure 2.

(1) Method using orthophoto

In this method, single photograph is oriented using several ground control points first. Second, grid with specified resolution is set on the digital map. For each grid point, corresponding photo coordinate is calculated from orientation parameter and digital elevation model (DEM), and pixel value is resampled by nearest neighbor or bilinear method. The generation of digital orthophoto in this way takes much time, but this method is very natural method.

(2) Method using affine transformed photograph

In this method, more than three control points which are distinct both on digital map and aerial photograph are selected first. Second, affine transformation is determined from

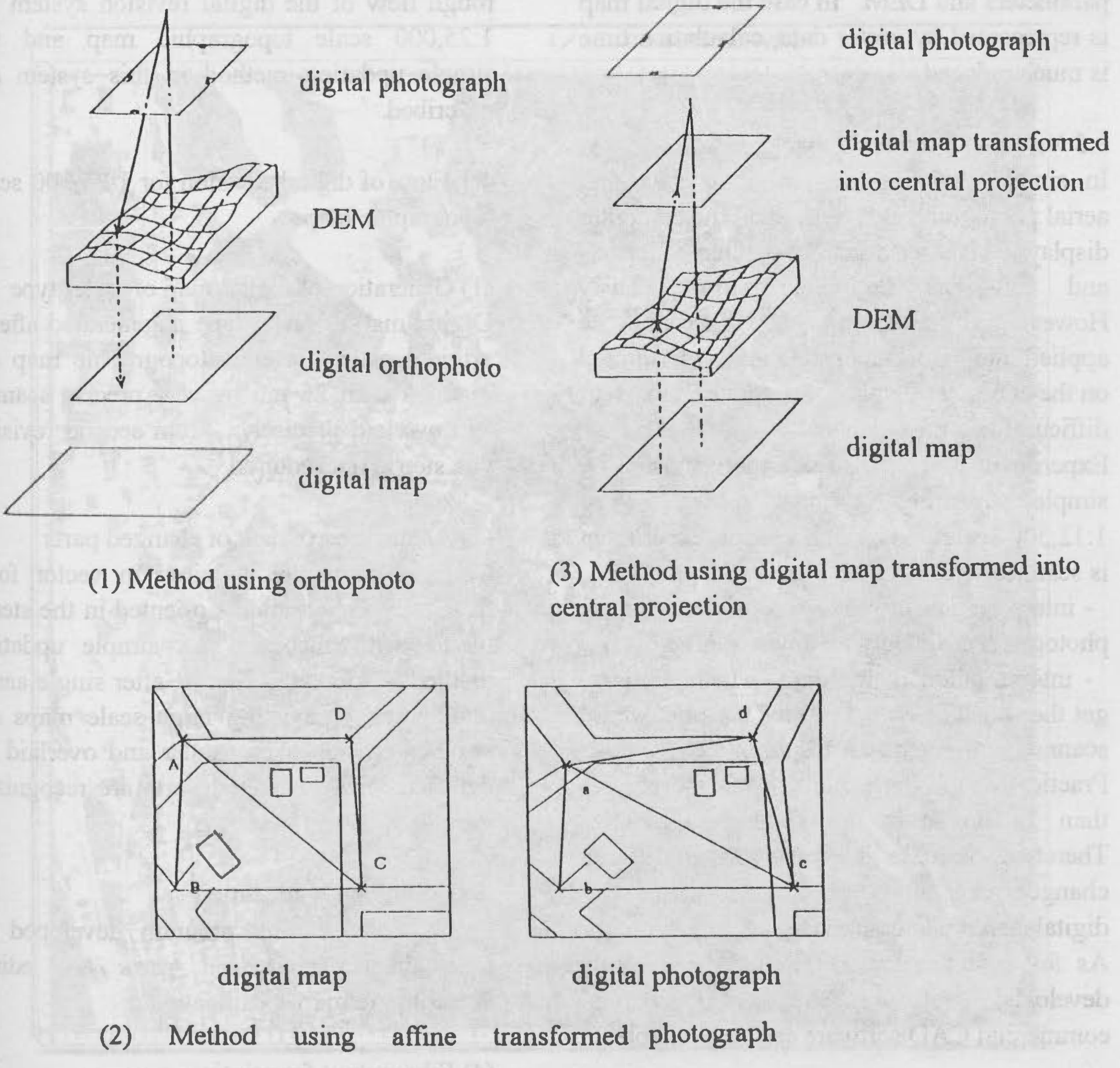


Figure 2 : Methods for overlaying photograph on map

control points by least square method. Finally, aerial photograph transformed by the affine transformation is overlaid on digital map. This method is not suitable for hilly or mountainous area due to local distortion by topographical effect, but this is effective for flat area which include most of urban area and does not require DEM.

(3) Method using digital map transformed into central projection
After single photograph orientation is conducted, digital map data is transformed on aerial photograph using orientation parameters and DEM. In case the digital map is represented by vector data, calculation time is much reduced.

3.3 Visual extraction of changed parts and edit
In simple updating system, map data and aerial photograph are overlaid on the computer display, visual extraction of changed parts and edit can be done simultaneously. However, as stereoscopic viewing cannot be applied and visual interpretation is performed on the computer display, interpretation is very difficult for some objects.
Experiment of 1:2,500 scale map revision by simple updating system shows, when 1:12,500 scale photograph used in experiment is scanned with from 12.5 μm to 25 μm pitch,
- interpretation of roads can be done for the photograph scanned with any scanning pitch
- interpretation of building, which needs to get the shape of each, requires the photograph scanned with finer than 15 μm pitch.

Practically, scanning pitch should be coarser than 25 μm from the point of data size. Therefore, in case interpretation is difficult, changed parts should be extracted using non-digital stereo photographs.
As for software for edit, not only specially developed software but also several commercial CAD software can be available.

4. UPDATING SYSTEM FOR 1:25,000 SCALE TOPOGRAPHIC MAP

Our institute, GSI, is responsible for maintaining 1:25,000 scale topographic maps which are largest scale maps covering whole Japan. In order to conduct revision work of these maps smoothly, digital revision method has been developed and is supposed to be adopted as the main revision method (Koide et al.,1992). As the method for getting changed parts, the simple updating method is newly included, adding to the ordinary stereo method using stereoplotter. In the following, rough flow of the digital revision system for 1:25,000 scale topographic map and the simple updating method in this system are described.

4.1 Flow of digital revision for 1:25,000 scale topographic maps

(1) Generation of digital map of raster type
Digital map of raster type is generated after 8 original plates for each topographic map are scanned with 25 μm by very precise scanner and overlaid precisely. From second revision, this step is not required.

(2) Acquisition of data of changed parts
Changed parts are acquired in vector form from stereo photographs oriented in the stereo plotter with encoder. Or simple updating method is adopted. That is, after single aerial photograph or existing large scale maps are scanned by A4 size scanner and overlaid on the raster map, changed parts are recognized visually.

(3) Editing on workstation
Using raster editing program developed by GSI, data of changed parts are edited according to map specification.

(4) Film output for printing

Revised raster map is output onto film by laser plotter.

4.2 Merit of the simple updating system for 1:25,000 scale topographic map

This system is devised to be used practically for map revision. First of all, considering aerial photograph used in the present revision work is not film but print and A3 size scanner, which can scan full size of aerial photograph, is expensive, A4 size scanner is adopted. Second, scanning pitch is more than 60 μm (up to 100 μm) and scanned photograph is represented with 32 colors (or grey scale), which makes data volume not so large. Third, reduced image from 1/2 to 1/32 are kept in order to make the operation smooth.

Besides these devices, as generalization is often conducted in the editing work of 1:25,000 scale topographic map, precise shape of each building is not required, which makes the demerit of simple updating method very small. As the demerit of this system, orientation in mountainous area is difficult because control points are not acquired easily. Figure 3 shows the example of aerial photograph overlaid on raster map.

Figure 3 : Example of aerial photograph overlaid on raster map

5. UPDATING SYSTEM FOR 1:2,500 SCALE TOPOGRAPHIC MAP

As large scale mapping is widely demanded, the influence of practical simple updating system is expected very big. So far, prototype systems are built up.

5.1 Characteristics of simple updating system for large scale map

In large scale mapping, as the shape of each building should be drawn, small structures in the digitized photograph are required to be visually clear. Therefore, in this system photograph is digitized in high resolution with very accurate scanner. As a result, data size of digital photograph become very huge, which means high performance personal computer or engineering workstation is required. In addition, digital map data are generally vector data, so "method using digital map transformed into central projection" is used for overlay.

5.2 Experiment of simple updating using prototype system

Using prototype system, experiment of simple updating and accuracy assessment have been conducted. The digital map revised by prototype system was compared with the map revised by ordinary 3 dimensional measurement using stereoplotter. Accuracy of simple updating method was assessed by root mean square of the difference of coordinates of corresponding points in these two revised maps. Accuracy is calculated for both roads and buildings respectively. Table 1 shows specification of data used for experiment and the accuracy of simple updating method. Figure 4 shows revised map by simple updating method (solid lines) overlaid on that by 3 dimensional measurement (broken lines).

6. CONCLUSION

Simple updating system for digital map, which does not require special photogrammetric instruments but common personal computer or workstation and anyone can handle easily, has been developed.
Especially, system for 1:25,000 scale topographic map which has developed by GSI is very practical and is supposed to be adopted in the routine revision work from this year. Prototype systems for large scale mapping have also developed. Those systems show fairy good accuracy, but visual interpretation of fine structures on digital photograph such as shapes of buildings is difficult unless photograph is digitized in very high resolution.

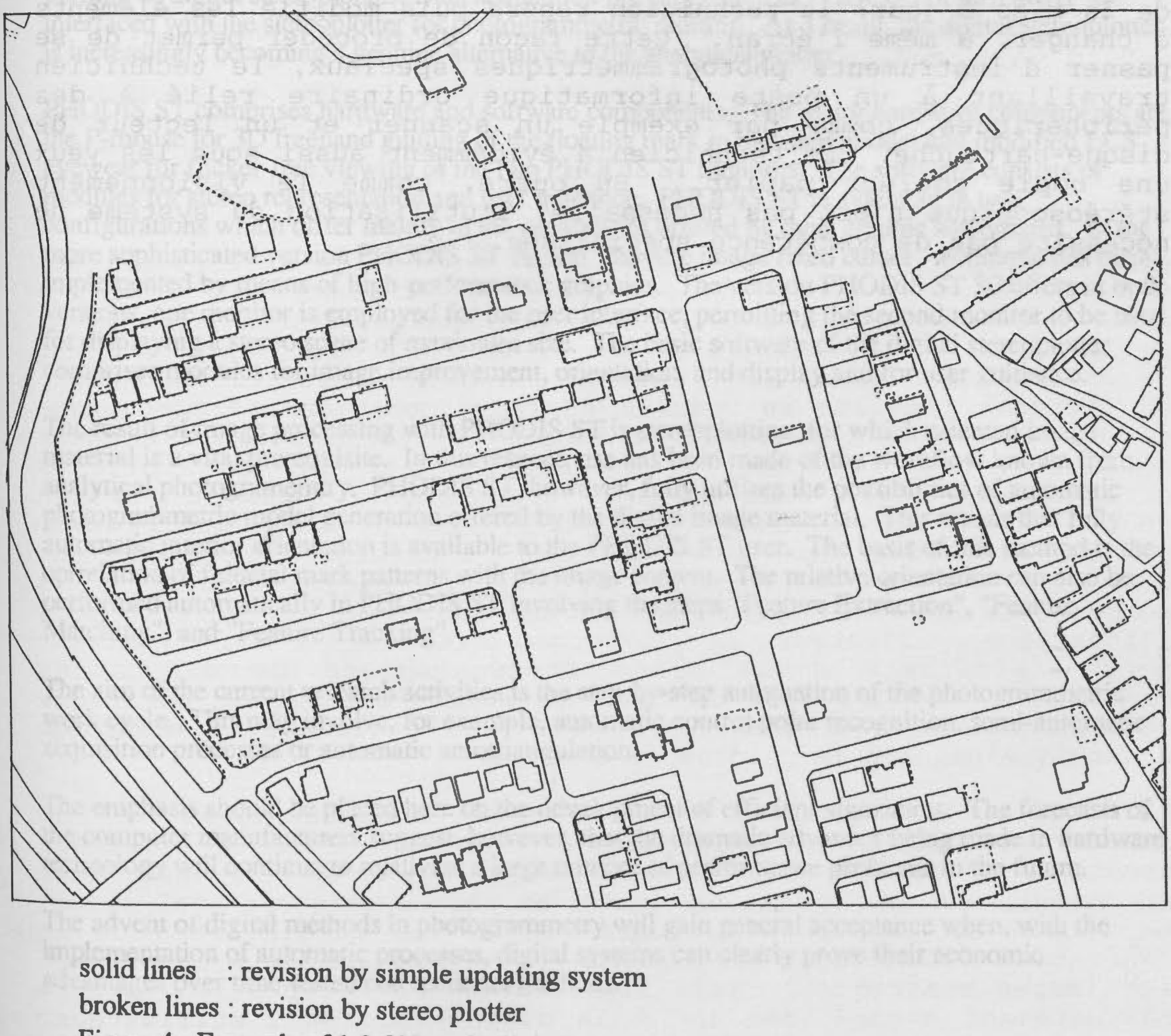
REFERENCES

Koide, M., I. Kamiya, T. Yoshinari, and B. Urabe, 1992. Raster based 1/25,000 Topographic Map Revision. International Archives of Photogrammetry and Remote Sensing Commission IV : 517-522.

digital map data to be revised	1:2,500 scale digital map data
scale of aerial photograph	1:12,500
scanning pitch	11.5 um x 13.5 um
accuracy of roads	0.57 m (52)
accuracy of buildings	1.23 m (28)
rate of omission for building revision	7 %

- planimetric accuracy of the revision by simple updating method assuming revison results by stereo plotter true
- figures in parenthesis are numbers of points where differences are measured

Table 1 Specification of data used for experiment and the accuracy of simple updating method



solid lines : revision by simple updating system
broken lines : revision by stereo plotter

Figure 4 : Example of 1:2,500 scale digital map revised by simple updating system

Mise au point d'un système simple de mise à jour des cartes numériques

RÉSUMÉ

Les cartes numériques constituent une information de base essentielle des SIG; il est donc indispensable qu'elles soient fréquemment mises à jour. Cependant, la méthode traditionnelle de mise à jour, au moyen d'un appareil de restitution photogrammétrique et de stéréophotographies, requiert le concours de spécialistes et impose à ceux-ci un travail long et fastidieux.

C'est pourquoi, en collaboration avec d'autres entreprises, nous avons créé un système de mise à jour pouvant simplifier énormément tout ce travail. Les photographies aériennes numériques sont d'abord superposées à une carte numérique, en vue de la mise à jour; le technicien repère puis modifie les éléments à changer, à même l'écran. Cette façon de procéder permet de se passer d'instruments photogrammétriques spéciaux, le technicien travaillant à un poste informatique ordinaire relié à des périphériques, comme par exemple un scanner et un lecteur de disque-cartouche. Le technicien a évidemment aussi sous les yeux une copie couleur papier. En outre, comme le visionnement stéréoscopique n'est pas nécessaire, l'utilisation du système ne nécessite pas de compétence spécialisée.

PHODIS ST - Design and Integration of Carl Zeiss Digital Stereoplotter

Phillip Wilkomm, Christoph Deerstel, Wemer Mayr
Carl Zeiss, Germany

ABSTRACT

Photogrammetry is currently undergoing a dramatic change from analytical to digital techniques. For the manufacturer of a photogrammetric workstation, this necessitates a critical assessment of both worlds, with the aim of optimally adapting the new methods to practical requirements. Carl Zeiss has entered this new era with the launching of the PHODIS digital photogrammetric system.

The PHODIS system comprises all components required in digital photogrammetry - from image generation and image processing right up to image output. The latest module in this package is the PHODIS ST digital stereoplotter. PHODIS ST is used for the generation, visualization and plotting of digital stereomodels. Data acquisition packages such as PHOCUS or CADMAP are interfaced with the stereoplotter for photogrammetric plotting. As a result, the digital stereoplotter is increasingly becoming a genuine alternative to the analytical plotter.

PHODIS ST comprises hardware and software components. The major hardware components are the P-mouse for 3D freehand guiding of the floating mark in the stereomodel and modified LCS eyewear for flicker-free viewing of the two PHODIS ST monitors. The software consists of modules for stereo representation and user guidance. PHODIS ST is available in two configurations which differ mainly in the capabilities offered by their graphic subsystems. In the more sophisticated version PHODIS ST 10, the "moving image fixed cursor" technique has been implemented by means of high-performance graphics. The version PHODIS ST 30 offers in both versions, one monitor is employed for the user interface, permitting the second monitor to be used for displaying a stereo scene of maximum size. The basic software of the digital stereoplotter comprises modules for image improvement, orientation, and display and for user guidance.

The result of image processing with PHODIS ST is stereoplotting, for which oriented image material is a vital prerequisite. In this respect, use has been made of the workflow known from analytical photogramemtry. PHODIS ST, however, fully utilizes the possibilities of automatic photogrammetric model generation offered by the digital image material. This means that fully automatic interior orientation is available to the PHODIS ST user. The basis of this method is the correlation of fiducial mark patterns with the image content. The relative orientation can also be performed automatically in PHODIS ST involving the steps "Feature Extraction", "Feature Matching", and "Feature Tracking".

The aim of the current research activities is the step-by-step automation of the photogrammetric work cycle. This may involve, for example, automatic control point recognition, semi-automatic acquisition processes or automatic aerotriangulation.

The emphasis should be placed here on the development of efficient algorithms. The forecasts of the computer manufacturers suggest, however, that the dramatic advances being made in hardware technology will continue to relativize a large number of performance problems in the future.

The advent of digital methods in photogrammetry will gain general acceptance when, with the implementation of automatic processes, digital systems can clearly prove their economic advantages over time-tested conventional methods.

PHODIS ST - Conception et intégration du stéréorestituteur numérique Carl Zeiss

RÉSUMÉ

La photogrammétrie subit actuellement des bouleversements saisissants pour passer des méthodes analytiques aux méthodes numériques. Pour le fabricant de postes de travail photogrammétriques, cela exige une évaluation critique des deux domaines visant à adapter de manière optimale les nouvelles méthodes aux exigences pratiques. Carl Zeiss aborde cette ère nouvelle par le lancement du nouveau système numérique de photogrammétrie PHODIS.

Le système PHODIS comprend toutes les composantes nécessaires en photogrammétrie numérique - depuis la génération et le traitement d'images jusqu'à la restitution d'images. Le plus récent module de cet ensemble est le stéréorestituteur numérique PHODIS ST. Le PHODIS ST sert à la génération, à la visualisation et au tracé de stéréomodèles numériques. Des progiciels de saisie de données comme le PHOCUS ou le CADMAP sont interfacés au stéréorestituteur pour la restitution photogrammétrique, ce qui fait que le stéréorestituteur numérique devient de plus en plus une véritable solution de remplacement du restituteur analytique.

Le PHODIS ST comprend des composantes matérielles et logicielles. Les principales composantes matérielles sont une souris-P permettant un guidage à main levée tridimensionnel de l'index de pointé dans le stéréomodèle et une visière LCS modifiée permettant le visionnement sans scintillement des deux moniteurs PHODIS ST. Le logiciel consiste en modules de représentation stéréoscopique et de guidage de l'usager. LE PHODIS ST est disponible en deux configurations différant principalement par les possibilités offertes par leurs sous-systèmes graphiques. Dans la version PHODIS ST 10 plus évoluée une méthode dite d'«image mobile avec curseur fixe» est mise en oeuvre par l'entremise d'un traitement graphique à haute performance. Dans la version PHODIS ST 30 commune aux deux configurations un des moniteurs est utilisé pour l'interface utilisateur, ce qui permet d'utiliser l'autre moniteur pour l'affichage d'une scène stéréoscopique de dimensions maximales. Le logiciel de base du stéréorestituteur numérique comprend des modules d'amélioration, d'orientation et d'affichage d'images ainsi que de guidage de l'usager.

Le résultat du traitement d'images au moyen du PHODIS ST est une stéréorestitution pour laquelle l'utilisation d'images orientées constitue un condition préalable. À cet égard, l'on a fait appel au déroulement des opérations connu en photogrammétrie analytique. Le PHODIS ST exploite toutefois toutes les possibilités de la génération automatique de modèles photogrammétriques qu'offrent les images numériques. Cela signifie qu'une orientation interne parfaitement automatisée est à la disposition de l'utilisateur. Cette méthode est basée sur la mise en corrélation des configurations des repères de centrage avec le contenu des images.

Le PHODIS ST permet également d'effectuer automatiquement l'orientation relative par les étapes «extraction des caractéristiques», «mise en correspondance des caractéristiques» et «poursuite des caractéristiques».

Les actuelles activités de recherche visent une automatisation étape par étape du cycle de travail en photogrammétrie. Cela peut englober, par exemple, la reconnaissance automatique des points de contrôle, des processus semi-automatiques de saisie ou l'aérotriangulation automatique.

L'emphase devrait ici être placée sur la mise au point d'algorithmes efficaces. Les prévisions des fabricants d'ordinateurs suggèrent toutefois que les progrès technologiques saisissants accomplis en matière de matériel continueront à l'avenir à relativiser un grand nombre de problèmes de performance.

L'avènement des méthodes numériques en photogrammétrie sera généralement accepté lorsque la mise en oeuvre de processus automatiques permettra de démontrer clairement les avantages économiques des systèmes numériques par rapport aux méthodes classiques éprouvées.

LA BARRE À PARALLAXE VIDÉO-NUMÉRIQUE

Jean-Paul Agnard, chercheur et P.-A. Gagnon, professeur
Université Laval, Québec

Commission II, inter-commission II/III

MOTS-CLÉS: Photogrammétrie numérique, barre à parallaxe, mesure de hauteurs.

RÉSUMÉ

La création d'un logiciel de pointé stéréoscopique, à l'université Laval, remonte à 1987. Conçu dans un cadre pédagogique, le logiciel visait à rendre plus conviviale l'utilisation de la barre à parallaxe par les étudiants en géomatique. Suivant cette conception initiale, un "BIP" prévenait l'étudiant quand la marque flottante atteignait la même altitude que le point mesuré, d'où enseignement par la machine et développement de réflexes conditionnés. Depuis ce temps, toutes les énergies du groupe de recherche en photogrammétrie numérique furent accaparées par la mise au point d'un logiciel de plus grande envergure dont l'ambition était de remplacer les appareils de restitution photogrammétrique. Nous pouvons dire maintenant: mission accomplie, le logiciel de restitution numérique DVP étant distribué par LEICA dans le monde entier. Pourtant, si l'on regarde le chemin parcouru depuis 1987, on s'aperçoit qu'une étape importante avait été occultée, celle de la mise au point d'une barre à parallaxe numérique, plus élaborée que le logiciel simplifié de pointé stéréoscopique, mais beaucoup plus simple que le vidéorestituteur numérique. C'est cette étape qui est visée par la communication présente dans laquelle il est montré, entre autres, que la différence de mesures de hauteurs d'objets avec la barre numérique et le DVP, dont la précision est le demi-pixel, est de l'ordre du quart de pixel.

The Digital Video Parallax-Bar

ABSTRACT

Traditionally, the parallax-bar has been part of basic courses in photogrammetry or photo interpretation. The first encounter of students with the tool, with its accompanying problems of photograph and micrometer orientations, has never been the type of experience designed to provide a good first impression of photogrammetry. This situation led to the idea of creating a software, to replace the traditionnal parallax-bar measurements. The first version (1987) dealt with simple figures displayed on the screen and on which the students had to make 3D pointings with a moving floating mark. A "BEEP" would be heard when the floating mark reached the correct elevation. This provided automatic training and development of conditioned reflexes. Based on the experience gained with the DVP (Digital Video Plotter) development, to which the research team since then directed most of its efforts, it was felt that the original idea had a very interesting potential of applications. This has led to the actual digital video parallax-bar software, which can be used by any one interested in measuring simple object heights. Experiments have shown that the difference between the DVP, the accuracy of which is half of a pixel, and the digital video parallax-bar, is of the order of a quarter of pixel.

1. INTRODUCTION

L'approche numérique a envahi presque tous les champs d'activités de la photogrammétrie. Qu'il s'agisse simplement d'entraînement au pointé tridimensionnel (Agnard et Gagnon, 1988), de restituteur (Agnard et al, 1988), de redresseur (Boulianne et al, 1992), d'aérotriangulation (Agnard et al, 1992), ou d'orthophotoplans (Novak, 1992), on peut presque trouver, actuellement, un équivalent numérique à toute méthode ou appareillage ayant existé de manière analogique durant des décennies.

Il ne reste que quelques rares exceptions. Certaines parce que leur mise au point est laborieuse, telle la caméra aérienne numérique, d'autres présentant moins d'intérêt et surtout moins de gloire. La barre à parallaxe numérique en est un bon exemple. Imaginer un tel instrument pourra même paraître pour certains quelque peu rétrograde ou du moins anachronique. Or, si l'on compare le nombre de barres à parallaxe actuellement sur le marché, on peut dire, sans se tromper, qu'il est de très loin supérieur au nombre de stéréorestituteurs existants. Dans l'optique qu'un jour, peut-être moins éloigné que l'on pense, les photographies aériennes seront toutes prises de manière numérique, et quand on pense qu'actuellement le prix d'un numériseur 8½"x11" à 300 DPI se compte en centaines de dollars, il n'est pas si ridicule qu'il pourrait le paraître à première vue d'envisager de faire également entrer la barre à parallaxe dans le monde de la photogrammétrie numérique.

Quand on sait, pour l'avoir utilisée, que la manipulation de la barre à parallaxe sous observation stéréoscopique d'un stéréoscope à miroirs demande une certaine dextérité, on comprendra aisément que la simple pensée de pouvoir effectuer les mêmes mesures, en regardant l'écran d'un moniteur de PC tout en tapant sur les touches d'un clavier, pourra rendre le sourire à plus d'un utilisateur de cet instrument de torture.

2. LA BARRE À PARALLAXE CONVENTIONNELLE

Pour mieux comprendre le côté laborieux de l'utilisation de la

barre à parallaxe, résumons-en les différentes étapes préparatoires.
Avant de pouvoir effectuer des mesures de parallaxe-x en vue d'en tirer des différences de hauteur, il faut tout d'abord orienter les deux photographies aériennes l'une par rapport à l'autre. Ceci consiste à aligner sur une même droite les deux centres de photos (à défaut des nadirs) et leurs homologues.
Cette première étape, assez facile à réaliser, franchie, il faut orienter le stéréoscope à miroirs (Figure 1) sur l'ensemble des

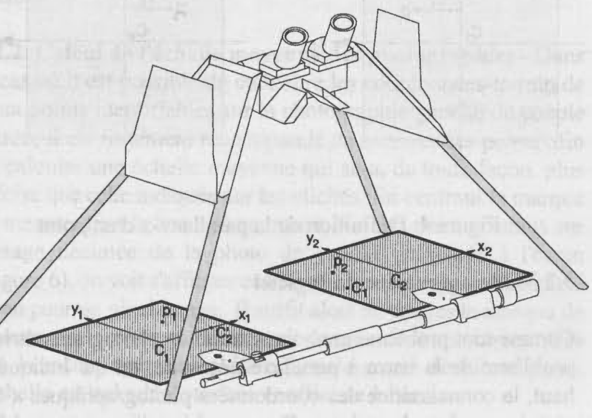


Figure 1: le stéréoscope à miroirs et la barre à parallaxe

deux photographies de façon à ce que les yeux soient parallèles à la ligne de vol définie précédemment. Ceci afin d'avoir une vision stéréoscopique confortable du terrain à mesurer.
Vient alors la partie la plus délicate de l'opération: orienter la barre à parallaxe sur les deux photographies, sous vision stéréoscopique, de manière à faire coïncider les deux marques de mesure (une sur chacune des deux plaquettes de verre de la barre à parallaxe) en une seule marque flottante. En tournant la vis micrométrique de la barre à parallaxe, cette marque paraît alors monter ou descendre dans l'espace virtuel en trois dimensions créé par le cerveau de l'observateur et non pas sous la table servant à la manipulation, comme l'ont si souvent enseigné faussement nombre de livres de référence sur la photogrammétrie lorsqu'ils traitent de la vision stéréocopique artificielle (Figure 2). Il ne faut pas que le défaut de coïncidence soit supérieur à une fraction du diamètre de

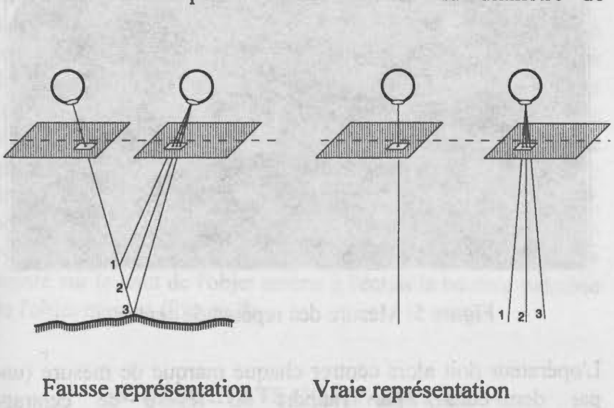


Figure 2: Fausse et vraie visions stéréoscopiques artificielles

chacune des deux marques, sinon il n'y a plus d'effet stéréoscopique et donc plus de marque flottante. Lorsque l'on essaie de faire cette coïncidence sans tourner le bouton moleté, cette manipulation est, avec un peu de pratique, assez simple. L'opération se complique lorsqu'il faut garder cette coïncidence tout en tournant le bouton moleté. Pour des néophytes, tels les 70 étudiants environ qui se présentent chaque année à l'université Laval aux cours de géomatique, il en résulte une grande perte d'énergie à tâcher de maintenir ces différents alignements, alors que la totalité des énergies dépensées devraient être entièrement consacrées au pointé stéréoscopique. C'est ce que nous avions fait ressortir dès 1988 au XVI congrès de l'ISPRS de Kyoto en y présentant notre logiciel d'entraînement au pointé tridimentionnel (Agnard et Gagnon, 1988).

3. LA BARRE À PARALLAXE NUMÉRIQUE

La photogrammétrie numérique nous a, depuis déjà quelques années, habitués à la numérisation des photographies aériennes sous forme de fichiers ou matrices de pixels; le pixel étant, pour une résolution de numérisation donnée, la plus petite quantité de découpage (la plupart du temps, carrée) de la photographie dont on enregistre le degré de gris, pour les photographies noir et blanc, compris entre le noir le plus absolu et le blanc le plus lumineux. Les numériseurs qui ne pouvaient donner, au début, que 16 niveaux de gris en ont fournis par la suite 64. Les numériseurs actuellement sur le marché en donnent couramment 256. Ceci représente au moins 200 de trop, puisque l'œil humain est incapable de distinguer plus d'une cinquantaine de teintes de gris (Gonzales et Wintz, 1987). Si ces 256 niveaux de gris sont intéressants pour, par exemple, calculer des corrélations d'images, il peut être avantageux de ramener, à l'affichage, ces 256 niveaux à 64, ce qui permet de consacrer les 192 autres niveaux pour des couleurs de vecteurs, texte, etc...

3.1 La mathématique impliquée dans la barre à parallaxe

Contrairement aux équations de colinéarité et (ou) de coplanéarité impliquées dans la mathématique utilisée dans les stéréorestituteurs analytiques, qu'ils soient conventionnels ou numériques, la seule équation impliquée dans la mathématique de la barre à parallaxe est des plus simples, pour la majorité des cas traités (terrain pas trop accidenté et petits angles de prise de vues ω,ϕ et κ):

$$\Delta h = \Delta p_x * (H-h_m) / p_{xb} \qquad (1)$$

où: **Δh** représente (Figure 3) la différence d'altitude entre les deux points de l'objet mesuré, c'est à dire sa hauteur,
Δp_x représente la différence des parallaxes-x des deux points mesurés,
$H-h_m$ représente la hauteur de vol au-dessus de l'altitude moyenne du terrain photographié,
et **p_{xb}** représente la parallaxe-x du point le plus bas des deux points mesurés.

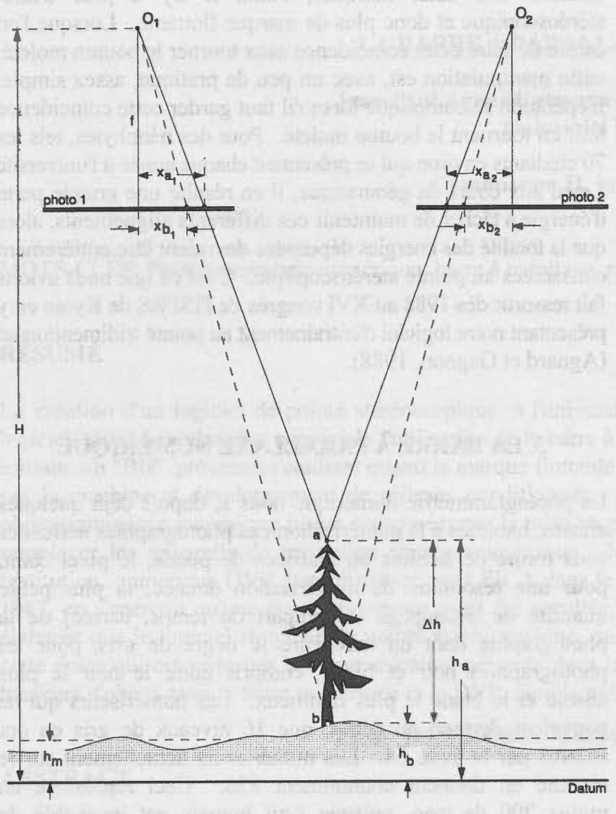


Figure 3: Mesure de la hauteur d'un objet

On sait, d'autre part, que:

$$H\text{-}h_m = f * E \qquad (2)$$

où: **f** représente la focale de la caméra de prise de vues,
et **E** représente le nombre échelle-moyen des deux photographies.

L'équation (1) devient alors:

$$\Delta h = \Delta p_x * f * E / p_{xb} \qquad (3)$$

Ce sont donc ces quatre éléments: Δp_x, f, E, et p_{xb} dont nous avons besoin pour calculer la hauteur d'un objet. Si la focale f de la caméra de prise de vues est une donnée, les trois autres éléments peuvent ou doivent être déterminés par mesures ou calculs à partir de manipulations sur les photographies. Le seul élément où le choix est possible est le nombre-échelle moyen des photographies qui peut, soit être entré directement à partir des informations existant sur les clichés, avec ce que cela suppose d'imprécision sur le chiffre indiqué, soit être calculé à partir d'au moins deux points connus en coordonnées sur le terrain.
Quant à p_x, la parallaxe-x, elle est définie (Figure 4) comme:

$$p_x = x_1 - x_2 \qquad (4)$$

x_1 et x_2 étant les abscisses de deux points homologues d'un même point terrain.

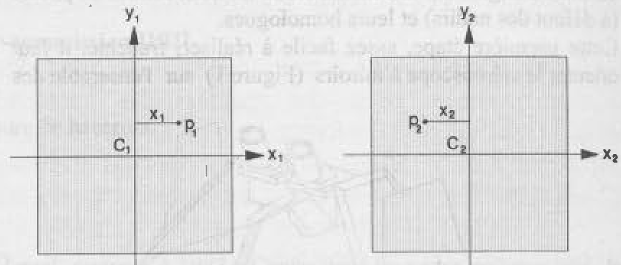


Figure 4: Définition de la parallaxe-x d'un point

3.2 Mode opératoire du logiciel

Comme tout problème analytique traité en photogrammétrie, le problème de la barre à parallaxe nécessite, tel qu' indiqué ci-haut, la connaissance des coordonnées photographiques x d'un certain nombre de points. Pour ce faire, l'image numérique étant représentée sous forme de pixels, il est nécessaire de faire une orientation intérieure, afin de pouvoir établir, pour chaque image, une transformation qui puisse faire le lien entre les coordonnées-pixels et les coordonnées photographiques.

3.2.1 Orientation intérieure - Après avoir indiqué au programme si les photographies ont été prises avec une caméra de type Wild (repères de centrage dans les coins) ou une caméra de type Zeiss (repères dans le milieu des côtés), il apparaît successivement à l'écran (Figure 5) les quatre couples gauche-droit de repères (un repère de centrage pour chaque demi-écran).

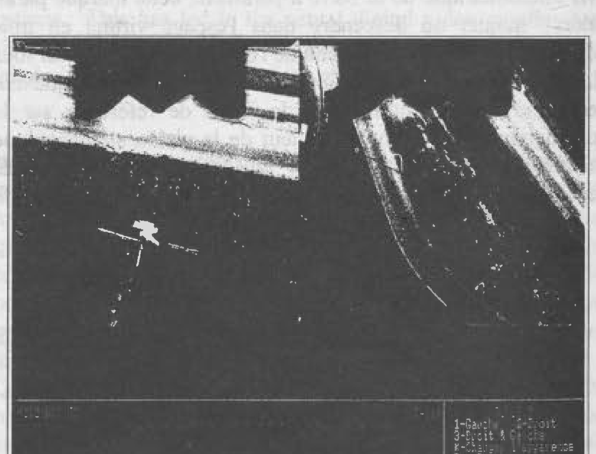


Figure 5: Mesure des repères de centrage

L'opérateur doit alors centrer chaque marque de mesure (une par demi-écran) par rapport au repère de centrage correspondant, à l'aide des flèches du clavier. Les deux marques de mesure qui, normalement, se déplacent de concert, peuvent être dissociées en appuyant sur la touche 1 du clavier (seulement déplacements de la marque gauche), ou sur la touche 2 (seulement déplacements de la marque droite). On peut à nouveau les rendre solidaires en appuyant sur la touche 3 (1 + 2). A la fin de la quatrième mesure, les résultats de l'orientation intérieure par transformation affine apparaissent à l'écran. Se trouvent ainsi corrigées le mieux possible, par moindres carrés, les déformations du film et celles du numériseur. Apparaissent alors à l'écran les coordonnées photo de chaque marque de mesure.

3.2.2. Calcul de l'échelle moyenne des photographies - Dans le cas où il est possible de connaître les coordonnées-terrain de deux points identifiables sur la photographie gauche du couple stéréo, il est fortement recommandé de mesurer ces points afin de calculer une échelle moyenne qui sera, de toute façon, plus précise que celle indiquée sur les clichés. En centrant la marque de mesure dans la zone où se trouve l'un des points connus sur l'image décimée de la photo de gauche présentée à l'écran (Figure 6), on voit s'afficher cette zone avec un rapport d'un pixel écran pour un pixel image. Il suffit alors de centrer la marque de mesure sur le point et d'entrer ses coordonnées au clavier. Dès que la même opération a été effectuée sur le deuxième point, l'échelle est calculée, stockée et affichée à l'écran.

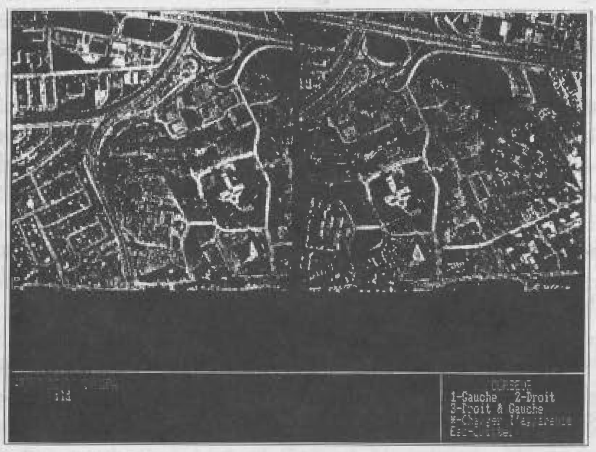


Figure 6: Les images décimées à l'écran

3.2.3. Mesure des abscisses et des parallaxes-x - Ces préparatifs étant terminés, l'opérateur est prêt faire ses mesures de hauteur d'objets. A ce moment, les deux images décimées étant de nouveau à l'écran , l'opérateur positionne la marque de mesure de gauche sur l'objet à mesurer dans l'image de gauche, puis celle de droite sur le même objet dans l'image de droite. L'effet de zoom précédemment expliqué se reproduit, permettant à l'opérateur de faire un pointé précis sur les deux points homologues du bas de l'objet à mesurer. Un autre pointé sur le haut de l'objet amène à l'écran la hauteur calculée de l'objet mesuré (Figure 7).

4. RÉSULTATS COMPARATIFS

Afin d'évaluer la précision de la méthode proposée, il nous est

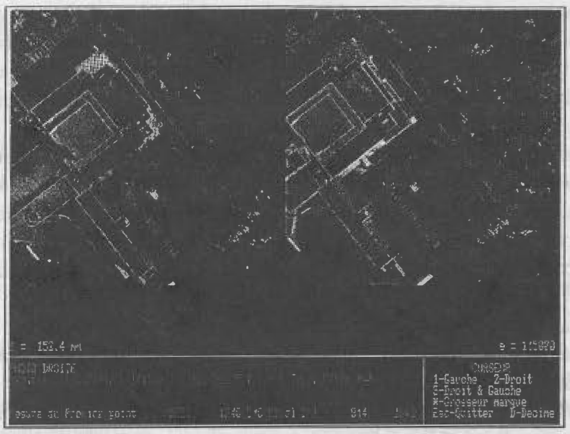


Figure 7: Mesure de la hauteur d'un bâtiment

paru logique de comparer des résultats obtenus à l'aide de la barre à parallaxe numérique avec les résultats obtenus au moyen d'un vidéorestituteur numérique. Ceci pour deux raisons. La première tient au fait que la comparaison peut être objective, du fait que dans les deux cas on peut se positionner sur les mêmes pixels, en haut et en bas de l'objet à mesurer. La deuxième parce qu'en comparant les résultats avec un système qui peut être à la fois de premier, second ou troisième ordre, suivant la finesse de la numérisation des photographies, 600 DPI, 300 DPI ou 100 DPI (nombre de points au pouce) on peut comparer la barre à parallaxe numérique avec un

Bâtiments	Δ h(m)barre	Δ h(m) DVP	Ecarts(m)
1	18.45	18.43	- 0.02
2	18.45	18.32	- 0.13
3	27.69	27.48	- 0.21
4	51.33	51.01	- 0.32
5	3.74	3.69	- 0.05
6	23.35	23.27	- 0.08
7	10.49	10.54	+0.05
8	4.51	4.58	+0.07
9	7.50	7.38	- 0.12
10	27.00	26.84	- 0.16

Tableau 1: Résultats comparatifs

système d'une précision supérieure à la barre à parallaxe conventionnelle.

Pour ce faire, les deux photographies numérisées utilisées pour les mesures à la barre à parallaxe numérique ont été orientées à l'aide du logiciel de photogrammétrie numérique DVP. La numérisation ayant été faite à 300 DPI et l'échelle moyenne des photographies étant de 1/5 000e, le pixel équivaut à 85 µm, soit à 42 cm sur le terrain. Ceci correspond à 71 cm en hauteur (42 * f / b, f étant la distance focale et b la base photographique). Dix hauteurs de bâtiments ont donc été mesurées avec les deux méthodes. Le tableau 1 nous montre ces résultats.

Les écarts obtenus sur les dix bâtiments, entre les deux méthodes, donnent un écart-type σ_h de 15 centimètres, soit 4.5 fois moins que l'équivalent en Z d'un déplacement en X d'un pixel. La précision du système DVP étant de l'ordre du demi-pixel (36 cm en Z dans ce cas), on constate qu'il n'y a aucune différence significative entre un vidéorestituteur numérique et une barre à parallaxe vidéo numérique en ce qui concerne la mesure de hauteur d'objets (le point haut et le point bas ayant sensiblement les mêmes coordonnées planimétriques).

Si l'on veut comparer ces résultats avec l'influence de l'erreur de lecture obtenue avec une barre à parallaxe conventionnelle, un grossissement du binoculaire de 3X et les mêmes photographies à l'échelle de 1/5 000e, on obtient, en dérivant la formule 1:

$$d\Delta h = d\Delta p_x * (H-h_m) / p_{xb} \qquad (5)$$

où: **$d\Delta h$** représente la précision sur la hauteur,
et **$d\Delta p_x$** la précision sur la différence de parallaxes,
en prenant pour acquis que les autres éléments de la formule sont, dans ce cas, des constantes.

La différence de parallaxes étant égale à la différence de lectures, on a:

$$d\Delta p_x = dl\sqrt{2} \qquad (6)$$

où **dl** est l'erreur sur la lecture du point haut ou du point bas.

Si l'on évalue cette erreur à 0.03 mm, on obtient une précision sur la hauteur mesurée égale à 36 cm. De la même manière que la précision obtenue avec la méthode conventionnelle est fonction du grossissement du binoculaire utilisé, la précision obtenue avec la méthode numérique est fonction de la résolution choisie pour effectuer la numérisation des photographies. Dans les deux cas, il existe des limites raisonnables à ne pas dépasser.

5. CONCLUSION

Le micro-ordinateur étant déjà présent dans tous les bureaux d'ingénieurs, de géographes, d'architectes, etc, il devient donc une barre à parallaxe numérique potentielle à la portée de tous ces utilisateurs éventuels.

Il est à espérer que la convivialité du logiciel associée à la facilité d'effectuer des mesures de hauteur, sans avoir à s'occuper de quelconques contraintes physiques, aidera à faire entrer la photogrammétrie numérique , sous forme de barre à parallaxe numérique, par la petite porte d'utilisateurs potentiels susceptibles d'utiliser un jour la photogrammétrie numérique sous une forme plus évoluée au potentiel illimité.

6. RÉFÉRENCES

Agnard, J.-P. et P.-A. Gagnon, 1988. L'enseignement du pointé stéréoscopique: une solution nouvelle à un problème permanent. In: International Archives of Photogrammetry and Remote Sensing, Kyoto, Japan, *******

Agnard, J.-P., P.-A. Gagnon and C. Nolette, 1988. Microcomputers and Photogrammetry. A New Tool: The Videoplotter. PE&RS, Vol. 54, No. 8, pp. 1165,1167.

Agnard, J.-P., C. Nolette and P.-A. Gagnon, 1992. PC-Based Integrated Digital Pugging and Measurements for Block Adjustment. In: International Archives of Photogrammetry and Remote Sensing, pp. 339-344

Boulianne, M., J.-P. Agnard et M. Côté, 1992. Redresseur d'images numériques. In: International Archives of Photogrammetry and Remote Sensing, pp. 160-165

Gonzales, R. C. and P. Wintz, 1987. Digital Image Processing, Addison-Wosley Publishing Company Inc., Second Edition, p.25.

Novak, N., 1992. Rectification of Digital Imagery. PE&RS, Vol. 58, No. 3, pp. 339-344.

NOVEL SOURCES OF CONTROL FOR AERIAL PHOTOGRAPHY

A.P.R. Cooper*, R. Swetnam** and A.J. Fox*

*British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge, United Kingdom.

**Department of Geography, The University of Edinburgh, Drummond Street, Edinburgh, United Kingdom.

KEY WORDS: Photogrammetry, Block adjustment, Aerotriangulation, Shape-from-shading

ABSTRACT

Sparse survey control data have constrained large-scale mapping of the Antarctic Peninsula. Extensive aerial photography is available but photogrammetric techniques are difficult to use because surveyed points are very widely spaced. The presence of extensive snow-covered areas with insufficient surface detail to allow the topography to be followed in a stereo-plotter creates a further difficulty. In this study, control information was taken from various sources including georeferenced satellite images, surveyed points and airborne barometric surveys. A shape-from-shading algorithm allowed the elevation of snow-fields to be interpolated from barometric height information. These data provided control for the block adjustment of a series of aerial photographs. A horizontal accuracy of ±60 m and a vertical accuracy of ±15 m was achieved in a block of photographs containing only one conventionally surveyed control point. Such errors allow the construction of an internally consistent large-scale map for an area where inadequate control had restricted mapping to medium and small scales.

INTRODUCTION

Sparse control data and the presence of snow-fields with insufficient surface texture for stereo-matching limit the usefulness of aerial photography in Antarctica. New control data are unlikely to be obtained because of the large areas to be covered, difficult access and severe climate. These difficulties have limited the scope for large-scale mapping even in regions well covered by aerial photography. In coastal or mountainous areas of Antarctica the largest scale maps with extensive coverage are at 1:200,000 or 1:250,000. To improve the quality of mapping in Antarctica, it is necessary to find means of using the aerial photography. This paper presents new techniques for creating sufficient control for a block of photographs that only contains one survey point. The study area is located in the north-east part of Colbert Mountains, Alexander Island, off the west coast of the Antarctic Peninsula (Fig. 1).

DATA

The data used in this study are from four sources:

1) The study area contains one surveyed point in the British Antarctic Survey (BAS) triangulation network. This triangulation network was adjusted to fit Geoceiver positions acquired in 1975/76 (Renner, 1982). The residual RMS error after this adjustment was ±11 m (Knight, 1986).

2) A Landsat TM image (scene ID 50719-12504 acquired on 18 February 1986) covers the area. It is one of a mosaic of images that was block adjusted to fit the BAS triangulation network by the Institut für Angewandte Geodäsie (IfAG), Frankfurt, using the method described in Sievers and others (1989). IfAG have corrected the

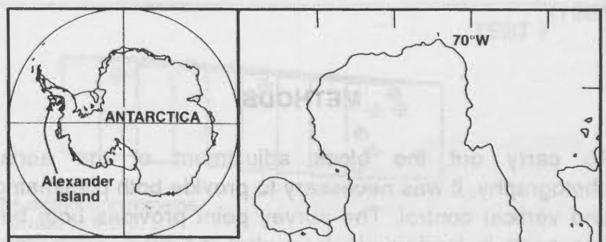


Figure 1: Location map for the study area.

image for the effects of surface elevation, using a coarse digital elevation model (DEM) based on elevation data from the Antarctic Digital Database (Thomson and Cooper, 1993).

3) A line of barometric surface elevations from a survey flight over an adjacent glacier were obtained in 1975 by BAS. These were measured by simultaneously recording the terrain clearance of the aircraft using a radar system and elevation from a pressure sensor. The pressure sensor readings were adjusted to height above sea-level by observing the terrain clearance when crossing features of known elevation, such as open sea or ice-shelf.

4) A block of aerial photographs. These were acquired on 1 January 1991, using a Zeiss RMKA15/23 camera mounted in a deHavilland Twin Otter. The flying height was 3800 m, resulting in a nominal plate scale of 1:25,000.

Additionally, satellite radar altimeter data were investigated. Measurements from the Geosat mission were available for the region, but none were found in the study area. If such data had been available, they would have been of great use, superseding the barometric surface elevations. Satellite altimeter data are only likely to be available for very smooth, flat regions such as snow-fields, because of limitations in the surface tracking systems on board the satellite (Rapley and others, 1983; McIntyre, 1991).

METHODS

To carry out the block adjustment of the aerial photography, it was necessary to provide both planimetric and vertical control. The survey point provides both but one point is inadequate to control the block. Two data sources were used to create extra control. In the first case, additional planimetric control was generated by using the georeferenced TM image. Points on the aerial photographs were identified on the satellite image, and their positions were determined using routines which were created within a GIS system.

Vertical control was more difficult to provide. The only additional source of height control was from the line of altimetric data from an airborne survey. Unfortunately, this was entirely over snow-fields, which presented insufficient detail on the aerial photographs for stereo-matching. A shape-from-shading technique was used to generate a digital elevation model of the snow fields from the satellite image. This shape-from-shading technique was developed by APRC, and only brief details are presented here. Further details are available on application to the senior author.

Shape-from-shading studies are based on numerical algorithms to solve (1), which describes the formation of an image from a surface with Lambertian scattering properties (Rouy and Tourin, 1992).

$$I(x,y)=\frac{\alpha\frac{\partial z}{\partial x}+\beta\frac{\partial z}{\partial y}+\gamma}{\sqrt{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2}} \qquad (1)$$

where I(x,y) is the image, α, β and γ are the components of the unit vector in the direction of illumination and x,y, and z are Cartesian coordinates with x and y in the horizontal plane and z vertical. For solar illumination, where the light source is very distant, the parameters α, β and γ are constant.

Equation 1 is difficult to solve, and under some circumstances may not have a unique solution. For many topographic surfaces, an important simplification can be made, i.e. for small slopes:

$$\sqrt{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2}\to 1 \qquad (2)$$

Substituting (2) into (1) gives:

$$I(x,y)=\alpha\frac{\partial z}{\partial x}+\beta\frac{\partial z}{\partial y}+\gamma \qquad (3)$$

As I(x,y) is arbitrarily scaled, the constant γ can also be ignored.

A solution to (3) can be found by taking the Fourier transform of the image. This results in the following relationships:

$$p_{n_1 n_2}=2\pi b_{n_1 n_2}\left(\alpha\frac{n_1}{N_1}+\beta\frac{n_2}{N_2}\right) \qquad (4)$$

$$q_{n_1 n_2}=-2\pi a_{n_1 n_2}\left(\alpha\frac{n_1}{N_1}+\beta\frac{n_2}{N_2}\right) \qquad (5)$$

where p and q are the coefficients of the Fourier transform of the image and a and b are the coefficients of the Fourier transform of corresponding elevations.

Thus, by taking the discrete Fourier transform of I(x,y) to obtain the coefficients p and q, the coefficients a and b can be obtained by applying equations 4 and 5. The surface is then constructed by taking the inverse Fourier transform of the resulting set of coefficients.

This technique has several advantages over direct integration of the image. The solution is a consistent surface for the entire area considered, with no discontinuities introduced by short wavelength features. Because the functions in (4) and (5) constitute a low pass filter in the frequency domain, the result is unaffected by high frequency features in the image, such as crevasses.

Test No.	No. of points	xy only	Errors		
			xy	z	xyz
1	6	0	37.6	15.3	30.9
2	14	5	176.4	30.4	140.0
3	8	0	60.5	13.6	47.5
4	13	5	259.2	201.8	193.0
5	9	0	59.6	14.6	47.1
6	7	2	178.1	30.1	140.0
7	10	1	170.7	30.8	136.7

Table I Summary of test results

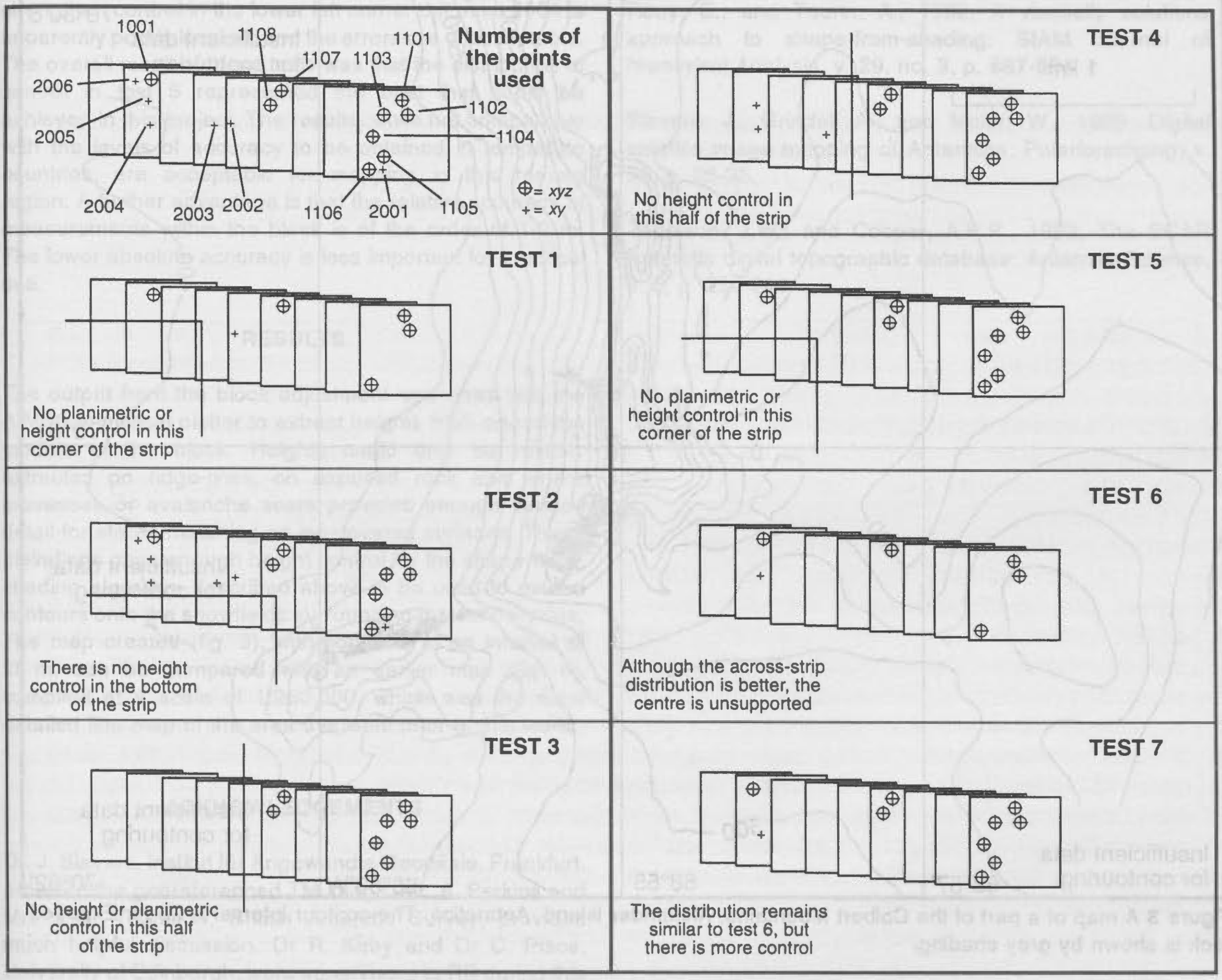


Figure 2: The distribution of control points used in various tests

This makes the technique insensitive to image noise and speckle. A final advantage is that the process is efficient, taking only a few minutes of computer time.

The elevation model created by this technique is arbitrarily scaled and may have a regional tilt. To provide surface elevations, it is necessary to calibrate the elevation model using measured surface heights. In this study, the model was calibrated using the airborne altimetric data, which allowed height values to be generated for regions of the snow-field where there was sufficient surface detail for successful stereo-matching.

CONTROL POINTS

The techniques and sources described above produced 15 control points, of which 9 had known elevation values. These control points are of three types: surveyed points (1 point), planimetric points from the TM scene (6 points) and heighted points from the shape-from-shading algorithm (8 points); the accuracy of the positions depends on their source. The surveyed points have a nominal accuracy of ±11 m (Knight, 1986), but the local geometry of the survey network greatly affects the accuracy of particular points. Because the image is tied to the same survey network as the surveyed points, the correspondence between the image and surveyed points is within one pixel (i.e. within 30 m) and the nominal accuracy with which points on the image can be located is ±30 m. As the TM scene covers a much larger extent than the study area, it was possible to verify the co-location of the image with survey points not used in this study. However, difficulties of identifying corresponding points on the TM image and the aerial photographs caused larger errors. Points on the edges of rock outcrops were located more reliably than those at the centre because of the high contrast between snow and rock. Errors may also arise from the different dates of the TM scene and the aerial photography. The extent of thin snow patches may alter substantially from season to season, and there is no means of quantifying this variation.

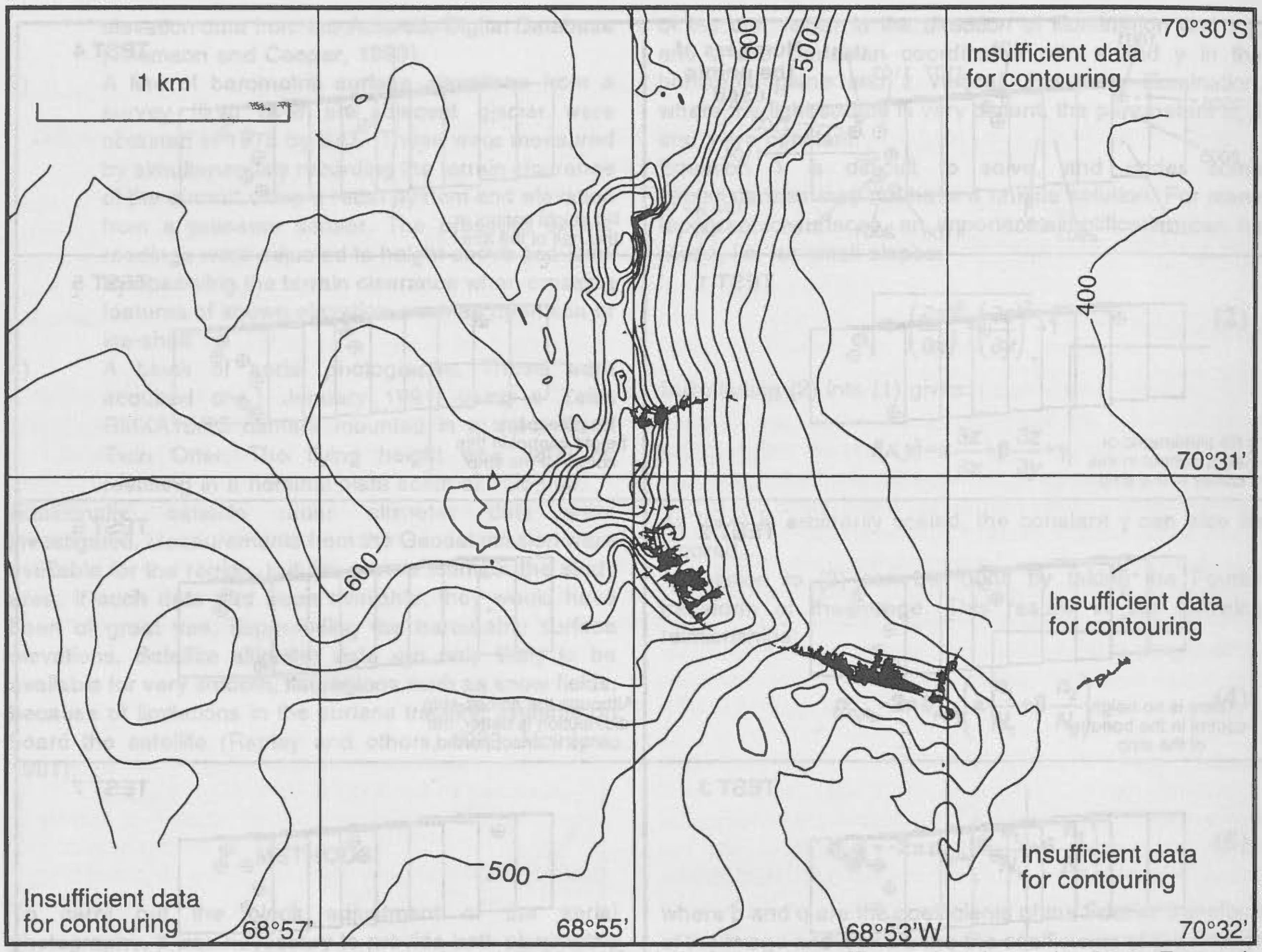


Figure 3 A map of a part of the Colbert Mountains, Alexander Island, Antarctica. The contour interval is 25 m. Exposed rock is shown by grey shading.

BLOCK ADJUSTMENT

The block adjustment was carried out using the BLOKK suite of programs provided with the CARTO Instruments AP190 analytical stereo-plotter. Several combinations of the control points were used to find the optimum selection. The best results have an RMS error of ±60 m horizontally,

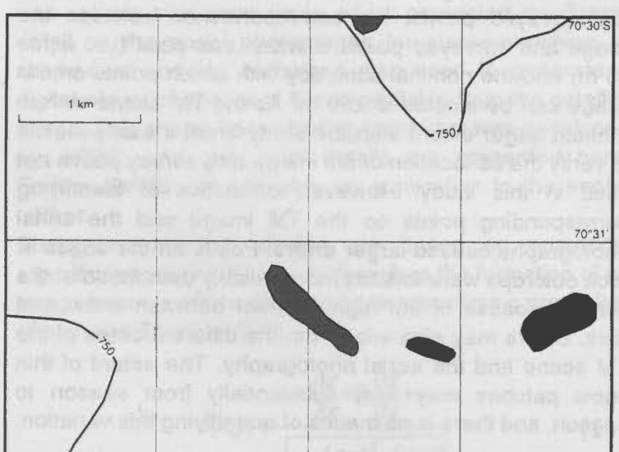


Figure 4 A map of the same region as fig. 3, compiled from satellite imagery at 1:250,000 scale in 1978.

and ±15 m vertically. The various combinations of points used are shown in Fig. 2, and the corresponding errors in Table 1.

The lowest errors were obtained from test 1, which only contained 6 control points. Although this is more than the minimum number required to control the block, the distribution of control points is weak, particularly in the lower left quadrant. Test 2 used all the available control and thus improved the distribution of points. Unfortunately, the error is much higher, suggesting that some of the control points are in error. Test 3 was designed to investigate the accuracy of the points derived from the shape-from-shading algorithm. This gave very good results but the distribution of control is unacceptable since the left-hand end of the block completely lacks control. However, the low error figures, especially in the vertical dimension, shows that the shape-from-shading algorithm has proved reliable. Test 4 attempted to add control to the left-hand side of the block and the poor result shows that some of the points in the lower left corner must have large errors. This is the only test where the vertical error is unacceptable. Test 5 used all the points with measurements in three dimensions, and the result is comparable with that from test 3; it gives a better distribution of control but the lower left corner is still unsupported. Tests 6 and 7 represent attempts to strengthen control in the lower left corner but point 2004 is apparently poorly located, and the errors are unacceptable. The overall result of these tests was that the distribution of control in test 5 represented the best that could be achieved in this project. The results, while not comparable with the levels of accuracy to be obtained in temperate countries, are acceptable for mapping in this remote region. A further advantage is that the relative accuracy of measurements within the block is of the order of 1-2 m. The lower absolute accuracy is less important for practical use.

RESULTS

The output from the block adjustment was used with the AP190 analytical plotter to extract heights from one of the models of the block. Heights could only be reliably extracted on ridge-lines, on exposed rock and where crevasses or avalanche scars provided enough surface detail for stereo-matching on ice-covered surfaces. These elevations gave enough height control for the shape-from-shading algorithm described above to be used to extend contours onto the snowfields surrounding the central ridge. The map created (fig. 3), with contours at an interval of 25 m, can be compared with an earlier map (fig. 4), compiled at a scale of 1:250,000, which was the most detailed line map of the area available prior to this work.

ACKNOWLEDGEMENTS

Dr. J. Sievers, Institut für Angewandte Geodäsie, Frankfurt, provided the georeferenced TM image. Mr. A. Perkins and Mrs J.W. Thomson, British Antarctic Survey, provided much helpful discussion. Dr R. Kirby and Dr C. Place, University of Edinburgh, were supervisors to RS during this project, which was undertaken as part of her MSc course.

BIBLIOGRAPHY

Knight, J.A., 1986, Report on the adjustment of the British Antarctic Survey Network: Cambridge, U.K., British Antarctic Survey; Internal report. Ref. No. ES2 EW 300/38, 36 pages [Unpublished].

McIntyre, N.F., 1991, Mapping ice sheets with the altimeter: International Journal of Remote Sensing, v. 12, p. 1775-1793.

Rapley, C.G., Griffiths, H.D., Squire, V.A., Lefebvre, M., Birks, A.R., Brenner, A.C., Brossier, C., Clifford, L.D., Cooper, A.P.R., Cowan, A.M., Drewry, D.J., Gorman, M.R., Huckle, H.E., Lamb, P.A., Martin, T.V, McIntyre, N.F., Milne, K., Novotny, E., Peckham, G.E., Schgounn, C., Scott, R.F., Thomas, R.H., and Vesecky, J.F., 1983, A study of satellite radar altimeter operation over ice-covered surfaces: ESA Report 5182/82/F/CG(SC). 224pages.

Renner, R.G.B., 1982, An improved gravity base-station network over the Antarctic Peninsula: British Antarctic Survey Bulletin, No. 51, p. 145-149.

Rouy, E., and Tourin, A., 1992, A viscosity solutions approach to shape-from-shading: SIAM Journal of Numerical Analysis, v. 29, no. 3, p. 867-884.

Sievers, J., Grindel, A., and Meier, W., 1989, Digital satellite image mapping of Antarctica: Polarforschung, v. 59, p. 25-33.

Thomson, J.W., and Cooper, A.P.R., 1993, The SCAR Antarctic digital topographic database: Antarctic Science, v. 5, p. 239-244.

Nouvelles sources de canevas pour la photographie aérienne

Mots-clés : Photogrammétrie, compensation par blocs, aérotriangulation, forme à partir des tons

Résumé

La rareté des données sur le canevas fait obstacle à la cartographie à grande échelle de la péninsule antarctique. On dispose de nombreuses photographies aériennes, mais les techniques photogrammétriques peuvent difficilement être appliquées car les points arpentés sont largement espacés La présence de vastes étendues couvertes de neige dont la surface ne présente pas suffisamment de détails pour que la topographie puisse être reconstituée par un stéréorestituteur soulève un autre problème. Dans la présente étude, les données relatives au canevas ont été tirées de différentes sources, notamment des images satellites géoréférencées, des points arpentés et des enregistreurs barométriques embarqués. Un algorithme servant à déterminer la forme à partir des tons permet d'interpoler l'altitude des champs de glace à partir des données de hauteur barométrique. Ces données fournissent le canevas nécessaire à la compensation des blocs d'une série de photographies aériennes. On a réussi à obtenir une précision horizontale de ± 60 mètres et une précision verticale de ± 15 mètres à l'intérieur d'un bloc de photographies ne comportant qu'un seul point arpenté selon les procédés traditionnels, ce qui constitue un progrès remarquable par rapport aux cartes établies à ce jour pour cette région et laisse présager la construction d'une carte à grande échelle qui possède la cohérence interne recherchée.

Evaluation of *spectral* classifiers for separating Sea Ice from open water in preparation for Radarsat.

T. Heacock[1], T. Hirose[1], M. Manore[2]

1. Noetix Research Inc. 9[th] Floor, 280 Albert St. Ottawa, Ontario, Canada, K1P 5G8.

2. Canada Centre for Remote Sensing (CCRS), Ice Applications Group, 588 Booth St, Ottawa, Ontario, Canada, K1A 0Y7.

ABSTRACT

The launch of Radarsat in 1995 confirms Canada's commitment to the use of spaceborne radar data for the monitoring of its land and ocean areas. Large volumes of data will be used by the sea ice community, and in particular the Ice Centre Environment Canada to monitor the shipping lanes in and around Canada. In an effort to assist in data analysis automated systems are being developed to extract value-added information products for end users. To date, research has been conducted on methods of extracting ice information from both calibrated and uncalibrated SAR data. The first step in the development of a fully automated system is the separation of areas of ice and water. Five algorithms that use only tone and local scene texture extracted from an uncalibrated SAR scene (so called *spectral* algorithms) were evaluated. These algorithms, representative of all *spectral* algorithms, were selected because of their computational speed and efficiency. The results illustrated that all the algorithms have the ability to separate ice from water in SAR imagery under ideal conditions. The performance varied on an image to image basis dependant on the amount of systematic or geophysical variability within each scene used for the evaluation. Observations about the relative strengths and weaknesses of the algorithms are made.

KEY WORDS

Synthetic Aperture Radar, Sea Ice, Classification, Algorithm, Tone, Texture.

1.0 INTRODUCTION

Remotely sensed data is a practical and cost-effective method of acquiring detailed, timely information over Canada's ice-infested waters. With the launch of RADARSAT in early 1995 (Langham 1992), large volumes of Synthetic Aperture Radar (SAR) data will be acquired over Canada's oceans. It is anticipated that manual analysis of the data on a regional scale will be overwhelming. To overcome this potential bottleneck, researchers have proposed the development of automated systems to generate products for end users in all applications disciplines, and in particular for sea ice applications because of the operational nature of the Ice Centre, Environment Canada (ICEC), and it's client services (Ramsay *et al.* 1993).

An automated system for classifying Radarsat imagery must be able operate on data acquired from a single channel, single polarized SAR (C-band hh). The system must be robust enough to operate on imagery illustrating a diversity of ice conditions, collected from all seasons and all geographic regions of operational interest to ICEC. Furthermore, the data product which will be used by ICEC operations will be ScanSAR data, generated through a quick delivery system. This data will contain image ambiguities that are a function of the unique nature of the ScanSAR image creation process.

To date, ice classification research has been carried out for both calibrated (Kwok et *al.* 1992) and uncalibrated (Barber et *al.* 1993 , Wackerman 1988), SAR systems. A calibrated SAR system allows a direct relationship between the pixel value and the backscatter of the target. If knowledge of the backscatter characteristic each ice type is known, the classification process simply assigns each pixel to the most similar ice type. However, this approach relies on each ice type having a unique and known radar signature, which is often not the case, particularly for new ice types. A calibrated classification system, for use with ERS-1 data (Cvv), is currently in place at the Alaska SAR Facility (ASF). It is anticipated that a calibrated system will be implemented at ICEC in the future for use with RADARSAT data (Chh). An uncalibrated SAR system, on the other hand also relates the pixel intensities to the backscatter from the target, but includes contributions from the radar system itself, and the variation of the signal within the scene and between scenes. The latter is a function of the backscatter response of the target, unequal weighting of the antenna pattern, and modulation within the radar scene.

The cornerstone of a fully automated ice classification system is the development of an ice/no ice classifier from which further value-added products, such as ice concentration can be generated. The fundamental component of the ice/no ice classifier will be the image segmentation algorithms. A human interpreter uses a variety of image attributes to segment an image, these include tone, texture, structure, shape, size and content (relationship between features). The tone, or pixel intensities, and texture properties are the lowest level elements that can be used to discriminate the major ice types within an image. Tone represents the backscatter (the amount of microwave energy reflected back to the sensor) while texture represents the spatial arrangement of the pixel intensities, which provides information on structure. The philosophy behind starting with elementary image properties is that the algorithms will be relatively simple and computationally fast, which is important to an operational implementation.d More complex properties can be added latter as needed.

In this paper the results of an evaluation of five algorithms that use pixel intensities and local texture to separate ice from open water in SAR imagery are presented. These algorithms, referred to as *spectral* classifiers, were representative of algorithms listed to date in the literature which use tone and local texture to segment a SAR image. Section 2 describes the datasets which were used to evaluate the algorithms. The datasets are described in terms of their unique geophysical and sensor characteristics. Section 3 describes the algorithms, and section 4 describes the processing results. In section 5 the causes of the algorithms successes and failures are explored, as well as what the anticipated ambiguities of Radarsat ScanSAR data will be and how they will affect the algorithms. The conclusions drawn from this research will be detailed in section 6. Finally, section 7 will comment on the future directions that need be considered.

2.0 DATASETS

Seven datasets were selected for evaluating the algorithms. These images provided a representative sample of the images that are analyzed by Ice Centre Operations. The characteristics of each of the datasets is described in table 1. The Ice Centre analyzes Star-2, X-band imagery (Falkingham, 1993), images numbered 3 to 7. In anticipation of Radarsat, a C-band image (image #1), collected by the CCRS *Convair 580* SAR system (Livingstone *et al.* 1987) was included for evaluation. Finally, for comparison between the C- and X-band data, an X-band image, (image #2), collected by the CCRS SAR coincident with image number 1 was also included in the evaluation. All images were horizontally polarized (transmitted and received).

The images illustrated tonal and textural variations within and between scenes,which are a function of the geographic location and season (i.e. ice type, ice surface wetness), and imaging characteristics (i.e. sensor frequency, range fall-off).

3.0 ALGORITHMS

A large number of *spectral* algorithms exist in the literature, each of which can be categorised by; 1) the discriminant function, 2) use of spatial statistics within a scene, and, 3) the strategy used to segment the scene. Five algorithms were selected for evaluation of their respective ability to separate ice from open water within an uncalibrated SAR image.

Table 1 Dataset Characteristics

	Location & date	Ice Conditions	Image Characteristics
1	Grand Banks East Coast, March 1989	• Predominant lead, surrounded by new, first year and brash ice. • Spring conditions, variably wet ice surface.	• Variation in target response with range for both ice and water. • Slight motion compensation error in near range. • Strong return (bright feature) from brash ice. • Weak return (dark feature) from wet ice surface. • Stronger return from open water in c-band than x-band imagery.
2	Grand Banks East Coast, March 1989	• Same as Image 1.	• Variation in target response with range for both ice and water. • SAR illumination problems, particularly in the near range. • Weak return (dark feature) from wet ice surface.
3	Beaufort Sea, August 1992	• Edge of arctic ice pack. • Ice Margin features; deteriorating wave broken firstyear and multiyear floes. • Pack ice features; conglomerate firstyear and mulityear pack ice, with distinct ridging and leads. • Flooded and thin ice visible within leads. • Strips of ice in open ocean.	• Strong return (bright feature) from multiyear ice, ridges and ice strips. • Low return (dark feature) from leads, open water and flooded wave broken ice. • Moderate return from new ice within the leads.
4	Gulf of St-Lawrence, February 1993	• Highly mobile and dynamic ice conditions during the winter ice advance. • Ice consist mostly of new ice formations; grease ice and nilas are visible. • Some grey-white ice (exhibiting some structure) is piled up behind madeleine island. • Heavily rubbled ice is visible in Northumberland Strait. • Open water is present behind Madeleine island and to the east of the ice edge.	• Scene consists of two Star-2 overpasses covering the entire Gulf region. • Fresh water ice (firstyear) from Saguenay river gave a weak return (dark signature). • New ice, consisting of sub-resolution pancake floes with raised edges generate a very strong (bright feature) return. • Open water, nilas and grease ice generate a very low return (dark return).
5	Barrow Strait High Arctic, August 1992	• Summer break-up and outflow of ice from Barrow Strait and adjoining channels into Lancaster Sound. • Large old floes are visible within broken/rubble firstyear ice. • Fastice is breaking out of the adjoining channels.	• Large first year ice floes are not distinguishable from large multiyear ice floes based on tone or texture alone. They are distinguishable based on shape due to surface wetness attenuating the radar return. • Open water has a weak return (dark feature). • Rubble ice has a very strong return (bright feature).
6	Jones Sound High Arctic, February 1990	• Consolidated mid-winter conditions. • Both multi-year, second year and thick first-year ice present. • Refrozen leads are visible within the multiyear ice. • Rubble fields are present.	• Excellent example of variations in return from same ice type; a second year floe appears mid-toned, heavily ridged multiyear ice appears very bright, while smooth multiyear ice appears dark. • Variation in signature is partly due to location in range, is partly a function of the source of the ice. • Refrozen leads and smooth firstyear ice have a low return (dark feature). • Rubble fields have a strong return (bright feature).
7	Labrador Sea, January 1992	• Advancing winter ice edge. • New ice, grey ice and strips and patches at edge. • Thin first-year or grey-white ice within the pack. • Open water visible as cracks within the pack, and along advancing edge. • Wave patterns visible with ice at margin.	• Radar return fall off in near and far range, centre of scene noticeably brighter. • Wave pattern noticeable by alternating strong and weak return.

The algorithms evaluated provided a good representative sample of spectral classifier algorithms. Table 2 summaries the properties of each of the algorithms evaluated. Detailed descriptions of each algorithm follow.

Algorithm A (Entropy) separates ice and water, based on local measurements of information randomness. A discriminate functions is used to describe the boundary between pixel representing ice and those representing water within scatterplots extracted from local windows passing over the entire image. The function which is selected for each window will have a negative slope and an intercept which is less than 128. This algorithm was proposed by Shokr et al. in 1991 and modified by Noetix Research. The modification was simply to use the mean pixel value and the standard deviation within the scatterplots, rather than using the pixel intensity and variance.

An assumption with this algorithm is that the backscatter from open water will always be lower than that of ice, and will have intensity values less than 128. When the boundary is set all pixels below it are classified open water, and all values above it are classified as ice.

Algorithm B (Migrating Means) separates classes based on a clustering technique which uses a histogram of the pixel intensities. In the first step, two randomly selected mean points are defined from the histogram. Two clusters are defined based on their proximity to the two mean points. Iterating redefines the mean points of each cluster and the data values are regrouped. This carries on until the clustering stabilizes.

Algorithm C (Polynomial) relies on determining the modality within a local window as it passes over the image. The modality is determined through the use of a fourth order polynomial fit which best describes the histogram for each local window. If the function is bimodal, the window is positioned over an area representing more than one class so it is ignored. If the function is uni-modal the mean of the local window is saved, and used as either an ice or water value. The separation of these means into ice and water is accomplished through the use of the migrating means theory.

Table 2 Algorithms

Algorithm	Name	Discriminant Function	Properties	Reference
A	Entropy	Linear	Samples from entire image	Shokr *et al.* (1991)
B	Migrating Means	Non-linear	Samples from entire image	Everitt (1974)
C	Polynomial	2nd Order, Linear	Samples from partitions	Wackerman *et al.* (1991)
D	Mask	Non-Linear	Initially, samples from entire image, then focuses on ice and open water areas	Noetix Research (1993)
E	Hierarchical Network	Non-Linear	Samples are local on the high resolution image and progressively increases to be global in successive coarser images.	Pietkainen (1981)

Algorithm D (Mask) is based on an extension of the migrating means theory. The migrating means algorithm was modified to recursively continue to separate classes, only stopping after a pre-determined maximum number of clusters are separated, or a predetermined number of pixel values per cluster is achieved.

Algorithm E (Hierarchical Network) segments an image using a parent-child linked pyramid structure. The approach is a layered arrangement of arrays in which each array is half the size of the array one level up, and the bottom level contains the image which is be classified. A parent-child relationship between each pixel value and it's position within the image is not fixed and may be redefined at each iteration. For each node in level l there is a 4x4 sub-array of 'candidate child' nodes at level l -1. The node itself is a member of four such sub-arrays for level l +1 nodes. On each iteration the node is linked to a single one of these four higher nodes between level candidate parent nodes. After each node is linked, there will be between 0-16 'legitimate' children. The non-parent links then define windows in the image, and ultimately the image segments.

4.0 RESULTS

All the algorithms evaluated used samples extracted from over the entire image to derive class boundaries. As a result, they all had difficulties with scenes that exhibited any systematic variation in radiometric values with range. The fall-off in the signal at the far range is due to a power loss and the change in backscatter from a target with imaging incidence angle. Both of these factors can be significant in airborne SAR systems, which have low imaging incidence angles (between 60°-85° for Star-2, and 50°-70° for the CV 580 scenes) but are expected to be less with satellite platforms such as Radarsat. However, Radarsat ScanSAR data which will be used by ICEC is unique in that it will be a mosaic of several imaging modes of the satellite covering a range of incidence angles (20°-49°). The concerns that this brings forth will be discussed latter in the text.

Table 3 provides an overview of the strengths and weaknesses of each of the algorithms as determined by evaluation with the datasets selected. In summary, all algorithms performed marginally in complex scenes that contained floes with a range of tones. For example, multiyear ice in the Barrow Strait (image #5) had moderate to dark signatures due to the presence of surface melt water attenuating the signal. In most cases (standard mid-winter) multiyear ice will have a bright signature due to the dominance of volume scattering. In these instances, the spectral algorithms misclassified the flooded ice floes as open water. The problems which were observed in the algorithms' handling of the datasets can be related to the inherent characteristics of the algorithms. These characteristics fell into primarily into two categories; 1) filtering, and, 2) population bias. The effects of these characteristics are described further on an algorithm by algorithm basis.

Algorithm A (entropy) generated easily interpreted output. Areas classified as ice or open water were homogenous regions unlike the output of many of the other classifiers. This characteristic is attributed to the filtering of the images prior to input into the Entropy classifier. In this evaluation filter sizes of 9 by 9 and 11 by 11 were selected for examination. These filter sizes were considered optimum for speed of processing and acceptable level of image noise. The reduction of noise by filtering, however, that leads to more homogenous classifications, is traded-off with a loss of spatial detail. This is particularly evident for small features such as fractures, ridges, waves in ice and ice edges. Linear features are expanded and irregularly shaped features are reduced in extent. Furthermore, any illumination or range fall-off problems with the data are accentuated. On the other hand, in some instances, this attribute favours on the side of the algorithm. For example, areas of wave broken and flooded multiyear and firstyear ice floes with a similar signature as the open water in image #3 were correctly classified as sea ice. This is a result of fine detailed features (rubble and ridges) which had a strong return being blurred and identified as ice. Similarly, in the C-band scene (image #1) the filtering was instrumental in properly classifying the open water by subduing some of the strong texture observed in the lead in image # 1.

This smoothing characteristic which result from filtering will have repercussions on any further value-added products, such as ice concentration estimates, which may be generated from this output.

Algorithm B (migrating means) achieved similar results to those observed from algorithm C (polynomial). The similarity is expected as the only difference between the two algorithms is the method by which they each extract samples from the image. Algorithm B uses all samples extracted from the imagery, while algorithm C uses only pure one-class samples. The input image and the classifier are identical for the two approaches.

The output from this algorithm was noisy due to the input of the raw image and not a mean filtered image file, however, the delineation of fine detailed features was captured. Moreover, the migrating means and its derivatives (algorithm C and D) overestimate the amount of open water in all the test scenes. This tendency is attributed to the bias of the algorithm toward the class with the smallest population, which in our cases was open water. The discrimination function is drawn towards the largest mode of the histogram, resulting in a large number of that mode's pixel being mis-classified .

Table 3

Algorithm	Failures	Successes
A (Entropy)	• Any range ambiguities accentuated. • Smooth first year & new ice classified as water. • Wet multiyear & first-year ice misclassified as water. • Smooth fast ice with low return signature classified as water.	• Open water lead in c-band image was properly classified. • Easily interpreted output.
B (Migrating) (Means)	• Any range ambiguities very evident. • Classifies new & thin first year ice as water. • Wet multiyear & first-year ice misclassified as water. • Noisy output.	• Handles variable signal return due to range fall off.
C (Polynomial)	• Smooth first year & new ice classified as water. • Noisy output.	• Better handling of range effects than algorithms A or B. • Open water generally well delineated.
D (Mask)	• Uncontrolled number of classes. • First year ice classified as water. • Lead in c-band image partly mis-classified as ice.	• Best handling of range effects. • Separates wet ice into distinct intermediate class.
E (Hierarchical) (Network)	• Processing intensive. • Negative results.	

This problem is evident in the output for images 2 and 4. In these scenes, the majority of the pixel were high values (i.e. 180-255 range) resulting in a discrimination function being drawn in the middle of the 'ice mode', and everything lower than this function was classified as open water.

It should be noted that the overestimation of open water was accentuated in scenes where a 'fall-off' in the radiometric values was observed in range, and where floes had a low to medium pixel intensity.

<u>Algorithm C</u>, (polynomial) filters the image to find pure samples of ice and open water from the original image. The filter is used for finding uni-modal samples of ice and water. When a uni-modal (one class) sample is found, the mean is saved. The dynamic range described by the saved "water means" and "ice means" are used in a migrating means procedure to generate a discriminant function for two-class separation. In this case the choice of filter size is critical as it will be directly related to the size of the features within the scene. In our testing, a filter size of 40 by 40 pixel was used. This filter size worked well in certain situations where the scene was not complex, yet failed more in scenes where few single class areas of that size could be found. With only a few means saved for each class the discriminant function was not accurately defined, and the resulting function was not based on true class information.

The typical output from this algorithm was much noisier when compared with algorithm B. It is anticipated that preprocessing the imagery with a noise reduction filter would result in more homogenous areas of ice and open water classes. However, this would aggravate the problem of finding regions of pure class samples.

<u>Algorithm D</u> (Mask) is identical to algorithm B except that it takes the processing one step further. Once two clusters have been established, each cluster is further subdivided into two, provided the total number of samples exceeds a threshold value.

In complex scenes containing floes with a variety of pixel intensities (tones) and textures, this algorithm captured mis-classified ice floes. All other algorithms misclassified the intermediate toned ice floes as open water. By further subdivision of classes the thin ice within leads in image number 3, and the wet-surfaced firstyear and multiyear ice in image number 5 were not incorrectly classified as open water. However, the division of the open water class into two classes was problematic in scenes where both subclasses were open water.

This algorithm provided a unique method for identifying ice features that had a range of tones. However, a method to determine when a class contains a mixture of ice and open water, as well as an improved mechanism to decide when a cluster should be subdivided would improve the algorithm.

The results of algorithm E (Hierarchical Network) for two class separation indicated promising results over the size of areas processed (256 by 256 pixel). When the number of classes was increased the results were less favourable. The subarea extracted from image number 3 included multi-year ice floes (bright tones), new ice (medium tones), and open water leads (dark tones). In this case the distinction between the open water in the leads and the new ice was not possible.

5.0 RADARSAT ISSUES

Several issues need to be addressed with respect to implementing *spectrally* based algorithms for the processing of Radarsat Scansar data. Of particular interest, in the context of this paper, are the characteristics of the backscatter coefficients for ice and water in c-band horizontally polarized SAR data. The 500 km swath of Radarsat ScanSAR data will range over incidence angles from 20°-50°. Figure 1 illustrates the radar backscatter as a function of incidence angle for both C- and X-band SAR data during the winter and summer (Onstott, 1992). The figure illustrates a large amount of variability in C-band signatures over the range of incidence angles that will be collected. The X-band data on the other hand has a relatively uniform return over those same incidence angles. For C-band imagery, the variation in target response with range for both ice and water is significant. For example, for calm water the winter time response at 20° is approximately -15dB, where as at 50° is -40dB; a variability 25dB. Compared with X-band where the variability in response is only 10dB (from -25dB to -35dB).

6.0 CONCLUSIONS

Spectral classifiers were selected because of their computational speed and efficiency, which is of great importance for operational implementation. The five algorithms that were selected for evaluation were representative of all classes of spectral algorithms described to date in the literature. Seven image scenes from different geographic regions, seasons and acquired by different sensors provided a broad dataset with which to evaluate the properties of each class of algorithm and define their strengths and limitations. These data allowed for the evaluation of systematic effects observed in uncalibrated SAR data, and geophysical features that confuse the signatures of ice and water.

The results indicate that all algorithms have the ability to separate ice from water under ideal conditions. The performance of each individual algorithm on an image to image basis was variable depending on the degree of systematic or geophysical variability within the image. In particular, variations in backscatter as functions of incidence angle, or ice type, combined with the characteristics of parameters within the algorithm, (such as filter size), all combined to generated the results that were achieved.

This work was an important first step in the development of an operational algorithm for the separation of ice from water within uncalibrated SAR data. The evaluations afforded a greater understanding of the issues that need be addressed. Several recommendations for further work need be examined, they include;

An approach to process subareas of the image and produce a seamless output is required to overcome the systematic variation in the signal with range.

All the tested algorithms extract samples from an image without regard to their location and generate one discriminant function which is applied to all data within that scene. However, when the imagery contains systematic variation in pixel intensities in range, a single discriminate function is inadequate.

Characterisation of the expected stability of the image backscatter (intensities) from scene to scene should be conducted for all SAR systems to be used by a classifier.

A priori knowledge on image characteristics which will be stable from scene to scene can be used by an algorithm to improve its classification accuracy. This applies to the range of intensities for each ice type as well as systematic radiometric anomalies introduced by the sensor and image formation processes.

Spatial and structural segmentation schemes should be considered to overcome the limitations of the *spectral* classifiers.

Complex scenes containing floes with dark tones or areas with smooth firstyear ice can be confused with open water using tone and texture alone. It is possible to differentiate some of these floes from open water by spectral algorithms, however there is no reliable means to identify which areas are open water or sea ice based solely upon the spectral information. Shape and size are two simple measures which can be used to identify many of these features.

Other simple textural measures should be evaluated for augmenting the *spectral* classification.

Textural algorithms which account for both the natural texture elements in an image and the speckle resultant from the coherent SAR system, may be used to characterize ice types and open water. The texture measures could be used in conjunction with shape and size measures to give an indication of internal structure.

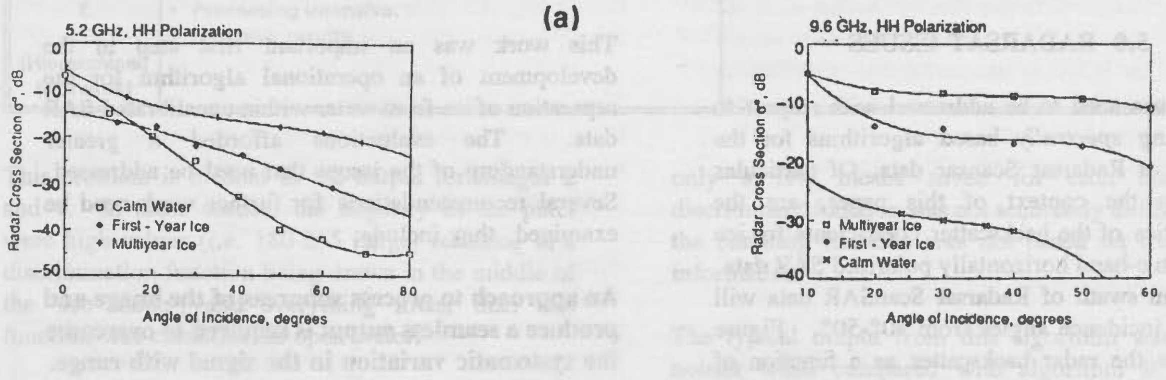


Figure 1 Radar backscatter cross sections at C and X band during winter (a) and summer (b), (Onstott 1992).

ACKNOWLEDGEMENTS

The authors kindly acknowledge valuable comments and input by R. Duncan. The authors would like to thank the Canada Centre for Remote Sensing, Department of Natural Resources Canada for the use of their facilities and to the Ice Centre Environment Canada for the SAR data. This work was supported through funding from the Radar Data Development Plan and the Panel on Energy Research and Development.

REFERENCES

Barber D., M. Shokr, R. Fernandes, E. Soulis, D. Flett, and E. Ledrew (1993) A Comparison of Second-Order Classifiers for SAR Sea Ice Discrimination. Photogrammetric Engineering & Remote Sensing, Vol. 59, No.9, pp. 1397-1408.

Everitt B., Cluster Analysis, published by Honeymoon Education Book, 1974.

Falkingham, J. (1993) Operational Remote Sensing of Sea Ice. Arctic, Vol. 44, Supp. 1, pp. 29-33.

Hirose T., (1992) User Requirements Specification for a Sea Ice Classification Algorithm using SAR Imagery Report prepared for the Ice Centre, Environment Canada, and the Canada Centre for Remote Sensing.

Hirose T., T. Heacock and R. Duncan (1993) A report on the evaluation of Ice Classification Algorithms. Report prepared for The Ice Centre Environment Canada, and The Canada Centre for Remote Sensing.

Kowk, R., E. Rignot, B. Holt, and R. Onstott (1992) Identification of sea ice types in spaceborne Synthetic Aperture Radar Data, Journal of Geophysical Research, Vol. 97, No. C2, pp. 2391-2402.

Langham E., (1992) Canada's Radarsat Program Proceedings of the 6th Australasian Remote Sensing Conference, November 1992, pp. 34-41

Livingstone, C., L. Gray, R. Hawkins and R. Olesen (1987) CCRS C-band Airborne Radar System Description and Test Results. 11th Canadian Symposium on Remote Sensing, Waterloo, Ontario, June 1987.

Onstott R. G., (1992) SAR Scatterometer Signatures of Sea Ice, Chapter 5, *Microwave Remote Sensing of Sea Ice*, Geophysical Monograph 68, pp 73-102.

Pietkäinen M., (1981) Segmentation by texture using pyramid node linking, IEEE T-SMC, Vol. 11, Dec. 81, pp. 822-825.

Ramsay, B., T. Hirose, M. Manore, J. Falkingham, R. Gale, D. Barber, M. Shokr, B. Danielowicz, B. Gorman, C. Livingstone (1993) Potential of Radarsat for Sea Ice Applications. Canadian Journal of Remote Sensing, Vol. 19, No. 4, pp. 353-362.

Shokr, M. and K. Zuberi (1992) Sea Ice Concentration from Synthetic Aperture Radar Images, The 12th Canadian Symposium on Remote Sensing, June 1992, pp. 368-373.

Wackerman *et al.*, (1988) Sea Ice Type Classification of SAR Imagery, Proc. of IGARSS '88 Symposium 1988, pp. 425-428.

Wells D. and D. Hagen, (1991) A knowledge Base to Support SAR Image Interpretation Systems for Arctic Ice Conditions, Phase 1: Pilot Project. A report prepared by Norland Science and Engineering for the Canada Centre for Remote Sensing.

[1]Évaluation de classificateurs spectraux servant à différencier la glace de mer et l'eau libre en vue de l'exploitation de RADARSAT

Les données de télédétection constituent un moyen pratique et rentable d'acquérir rapidement de l'information détaillée sur les eaux canadiennes envahies par les glaces. Après le lancement du satellite RADARSAT, prévu pour le début de 1995, on pourra recueillir des volumes considérables de données de radar d'ouvertureà synthètique (ROS) sur les océans du Canada. On prévoit que l'interprétation visuelle des données à l'échelle régionale deviendra une tâche extrêmement lourde. Pour éviter un éventuel engorgement, des chercheurs ont proposé de créer des systèmes automatisés qui serviraient à fabriquer des produits pour les utilisateurs finals de toutes les disciplines, et en particulier pour les applications des glaces, vu le caractère opérationnel du Centre des glaces d'Environnement Canada et des services qu'il offre à sa clientèle.

Pour analyser des images, les experts se fondent sur divers éléments d'information : la teinte, la texture, la structure, la forme, la taille et le contenu (rapports entre ces caractéristiques). Le système de classification automatique des glaces, quant à lui, aurait pour principale composante un ensemble d'algorithmes de segmentation d'images. La teinte et la texture sont des éléments d'information de premier ordre qui peuvent servir à différencier les types de glaces repérés sur les images ROS. La teinte fait référence à la rétrodiffusion (quantité de micro-ondes réfléchies de la glace de mer jusqu'au capteur), tandis que la texture représente l'arrangement spatial des densités de pixels, ce qui fournit des renseignements sur la structure. Les chercheurs ont choisi de travailler à partir d'éléments de premier ordre parce que les algorithmes utilisés seront relativement simples et rapides, et qu'il sera possible d'incorporer par la suite d'autres éléments plus complexes au besoin. Les classificateurs spectraux ont été retenus comme outils de départ parce qu'ils sont efficaces et rapides, ce qui constitue un critère important dans un contexte opérationnel.

Pour pouvoir fonctionner avec RADARSAT, le système de classification automatique devra être en mesure de traiter des données recueillies au moyen d'un ROS à canal et à polarisation uniques. La plupart des données ne seront pas étalonnées, et le système devra être assez robuste pour s'adapter à divers types de glace et pour traiter les ambiguïtés que l'on prévoit obtenir sur les images RADARSAT ScanSAR.

Les algorithmes spectraux possèdent trois caractéristiques : la fonction discriminante, l'utilisation faite des variations spatiales de la glace de mer dans une scène et la stratégie utilisée pour segmenter une scène. Le présent rapport porte principalement sur l'évaluation de cinq prototypes d'algorithmes spectraux de classification de la glace. Ces prototypes se fondent sur les variations de teinte et de texture locale pour différencier la glace et l'eau libre. Les auteurs se sont intéressés aux variations de teinte et de texture qui dépendent de l'emplacement géographique, de la saison (type de glace, humidité de la surface), des conditions climatiques et des conditions de prise d'image (fréquence du capteur et emplacement de la cible par rapport au capteur). Il faut absolument être en mesure de reconnaître ces effets pour pouvoir segmenter de façon exhaustive les images de glace de mer. On pourra ensuite se baser sur des images segmentées en glace et en eau libre pour approfondir la classification et offrir des produits à valeur ajoutée.

---

[1] Extended Abstract

PROCESSING ERS-1 SAR AND JERS-1 SAR FOR MAPPING

Ian Dowman and Mark Upton

University College London
Gower Street
London WC1E 6BT
United Kingdom

Tel: (44) 71 380 7226
Fax: (44) 71 380 0453

ABSTRACT

Synthetic Aperture Radar data from ERS-1 has been proved to be very stable and well calibrated and is of considerable interest for a number of geoscience applications. The production of JERS-1 data has been subject to a number of problems but is nevertheless of considerable interest. University College London (UCL) has used data from both sensors for geocoding and for the determination of heights using stereomatching. This paper reports on the characteristics of the data from both sensors and the processing required for geocoding using software developed at UCL and by a consortium in which UCL was involved. The derivation of the parameters needed for geocoding depended very much on the processor used for creating the images and was also dependent on the accuracy of the orbit data.

The determination of digital elevation models from both ERS-1 and JERS-1 also requires orbit data but depended on the radiometric characteristics of the data. The software derived for filtering and stereomatching the data is described and the results given.

The paper describes the problems encountered and the steps which have to be taken to achieve geocoding and stereomatching.

Keywords: SAR, ERS-1, JERS-1, geocoding, stereoscopy, speckle reduction filters, Digital Elevation Models.

1. INTRODUCTION

ERS-1 data has been available for some time now and the data has been used in the area of topographic mapping for geocoding and work has been done on the extraction of features. Results were presented at the first and second ERS-1 conferences, for example Dowman et al [1993b], Dowman et al [1994], Winter et al, [1993], Renouard and Perlant, [1994]. Work done at University College London (UCL) will be summarised and particularly the work on geocoding and stereomatching will be discussed.

The JERS-1 satellite with the L band SAR and an optical sensor (OPS) giving along track stereoscopic coverage adds important additional data to that already available for mapping. The SAR provides data at a different wave length and incidence angle to ERS-1 The data from JERS should therefore complement existing data from ERS-1 SAR and SPOT by providing SAR data which will highlight different surface features with the L band frequency, and allow two incidence angles with a potential for stereoscopic modelling for height derivation.

The paper starts with a report on geocoding of ERS-1 and JERS-1 and then gives a brief brief description of the method used for determination of height. The use of speckle reduction filters to improve the performance of stereo matching is then investigated. Some results are given which give an indication of the potential of the method and finally conclusions are drawn and the remaining problems are discussed.

2. PRINCIPLES OF GEOCODING SAR DATA

The principle of the system developed by the consortium in the UK (Dowman et al, 1993a) is the transformation of a SAR image in a co-ordinate system reflecting time and range data into a standard map projection. An image processed to zero Doppler is assumed. The transformation is carried out by mapping ground co-ordinates into the image by solving the equations relating the Doppler shift to the sensor and image co-ordinates taking into account the movement of the sensor and of the earth. The ground will be represented by a digital elevation model (DEM). This mapping is done exactly for a number of points arranged on a three dimensional framework, know as a supergrid, surrounding the area to be mapped, all other points are mapped by interpolation. The number of points rigorously mapped will depend on the nature of the terrain and the order of the polynomial chosen. The radiometric value to map back to the output image will be determined by resampling the image and a number of resampling routines will be available.

The method was developed and tested with ERS-1 data and the results are described in Dowman et al, [1993a] This paper will concentrate on results from JERS-1 SAR data.

3. GEOCODING RESULTS AND PROBLEMS FROM JERS-1 SAR

3.1 Test data

The region concentrated upon was the Marignane region of France. We have a DEM of this region, and have previously geocoded several ERS-1 images to it [Dowman et al 1993]. The DEM was generated by IGN France. The map projection is France Lambert zone 3, the sample spacing is 80m, and the sample quantization is 0.1m. The DEM covers an area approximately 16km x 16km, and elevations span the range [0m, 279m]. The geoid-ellipsoid separation for this region is in the range [7m, 7.4m], but as the DEM samples were assumed to be ellipsoid elevations, the separation was not applied. The projection from map to image was computed in the same map projection and on the same grid as the DEM, so no preprocessing was required.

The image was a JERS-1 SAR level 2.1 product [NASDA, 1992, Shimada, 1993], with a nominal pixel spacing of 12.5m. This is the standard SAR product. The map projection of this product is stated to be "geocoded" (as opposed to "ground range" or "slant range"), and polynomial transformations between GRS80 UTM zone 31 northern hemisphere and image coordinates are given. The polynomials are approximately rotations. Other documentation describes the coordinate system of the level 2.1 product as being azimuth and range. We therefore treat the image as being a ground range product. It should be noted that the product is alread geocoded only in the sense of ellipsoid correction; terrain correction is still necessary.

The sensor state vectors are given in ECR (Earth Centre Rotated), every 60 seconds. We have approximated ECR with WGS84. In the projection from global cartesian to slant range, the iteration termination threshold was set to 0.0001 seconds.

3.2 Tie pointing

The headers do not contain the information necessary to relate the azimuth coordinate of the image to azimuth time (zero-doppler or not). As a substitute, until a more satisfactory solution could be found, tiepointing was performed, as follows. Two points well separated in azimuth were chosen, and their azimuth coordinates measured in the image. Their coordinates were also measured in the map, then transformed to slant range, giving their zero-doppler azimuth times. A linear relation between the azimuth coordinate and the zero-doppler azimuth time was easily derived. This shows the time across one pixel to be approximately 0.0024 seconds, which justifies the choice of iteration termination threshold in the projection from global cartesian to slant range.

Samples of the range are given for the first, centre and last pixels in azimuth for each block of 1024 lines in range. They are only approximately zero-doppler, but were used as if they were exactly zero-doppler. The linear relation between the azimuth coordinate and the zero-doppler azimuth time, derived above, was used to generate azimuth time values on the same sample points. The resulting slant range samples were used to form the projection from slant range to ground range.

The final tiepoint correction to the overall projection requires that tiepoint coordinates be measured in ground range and in the image. We chose to do this step by reference to the DEM rather than published maps, as follows. First, the overall projection was derived without tiepoint correction. This projection was then used to generate an energy conservation map in image space, to be used as a ground range simulated image. The required tiepoint coordinates were then measured manually using an interactive image display tool in both the ground range simulated image and the input image, and used to form the projection from ground range to the image. The errors, above, in the approximations used to derive samples of the projection from ground range to slant range arecorrected for in this tiepointing step, along with any other errors.

It would have been just as possible to get the ground range tiepoint coordinates by projecting measurements from published maps. However, the approach taken avoided possible errors between the map and the DEM, and also allowed a greater number of tiepoints to be generated. It is usually very difficult to find a sufficient number of features that both appear and are well defined in both published maps and the image. A drawback of the approach taken is that the features used result from terrain effects, and so are less well defined in the azimuth direction than in the range direction.

After the overall projection was computed, the image was resampled to the map, completing the geocoding process. The ancillary products were also generated: In image space, layover and energy conservation; In map space, shadow, layover and energy conservation.

3.3 The geocoded result

In the range direction, the DEM slope is approximately in the range [-20°, 27°]. Shadow would be generated only at incidence angles greater than 70°, and layover only at incidence angles less than 27°. The scene centre incidence angle of JERS-1 is nominally 35°, therefore no shadow or layover from the DEM was to be expected. This is indeed the result which was obtained.

An important internal check on the geocoding is to compare the image space layover map against the saturated regions of the input image, and also to compare the map space shadow and layover maps against the saturated regions of he geocoded image. However, as these maps are (correctly) empty, all that can be done are comparisons between the energy conservation maps and the images. These show good geometric and radiometric agreement.

3.4 Problems with JERS SAR

There were many problems involved in determining from NASDA the exact definitions of the parameters in the headers. Some of these problems were solved late, some not at all. The consequence of this is that our geocoding process is still not properly matched to the standard JERS-1 SAR product. In this section, we discuss the remaining problems.

The coordinate system for the orbit is stated to be "ECR". "GSFC" is also mentioned in the same context. We have used WGS84 with no conversion. The actual relationship between these coordinate systems is not known.

We have treated the supplied range samples as if they were zero-doppler. It was subsequently confirmed that this is not the case. There has not been time to make appropriate modifications, but it is believed that the necessary parameters do exist in the headers. This does not invalidate the results above, as the tiepoint correction of the overall projection corrects for this.

We need to be able to relate the azimuth coordinate of the image to azimuth time. This is necessary so that use can be made of the orbit. We used a workaround involving two tiepoints and a linear approximation (described above) as a placeholder, hoping that NASDA would help us obtain the necessary information. Eventually, NASDA confirmed to us that the information does not exist in the standard product. It may exist in lower level products.

Amongst the various parameters in the headers of the standard product are corner coordinates of the image, given in geographic and map coordinates. In the product geocoded above, these were based upon the ellipsoid GRS80 and the map projection UTM zone 31 northern hemisphere. Projection of the corners to slant range would provide the necessary relation between the azimuth coordinate of the image and the azimuth time. However, investigation showed these corner coordinates to apparently be in error, preventing this approach. The error is large (1 - 4km), and cannot be accounted for by datum errors. NASDA could not help.

For the purpose of relating the azimuth coordinate of the image to azimuth time, the supplied scene centre time differs significantly (almost 600 pixels) from the value derived by tiepointing. This would, anyway, only have provided half of the necessary information.

The tiepoint correction of the overall projection from the map to the image can be performed using points measured in the map, or by using points measured in a "simulated" image, as done above. Both approaches have drawbacks. The use of published maps significantly limits the number of measurable tiepoint pairs, due to differences in content from that of the image. The use of an image simulated from the terrain gives poorer precision in the azimuth direction, due to the nature of features in such an image. Neither of these problems are specific to JERS-1 SAR; they are general to all SAR.

4. PRINCIPLES OF STEREO SAR FOR DEM PRODUCTION

The use of overlapping pairs of SAR images for the production of 3-D data has been outlined in Dowman et al (1992) and initial results presented in Dowman et al (1993b) . The process requires 3 stages: preprocessing of detected SAR images; stereo matching to determine disparities between the two images and transformation of the disparities to heights in a ground reference system.

The preprocessing to remove the effects of speckle is discussed in section 5 of this paper.

The method used for stereomatching is the CASCADE algorithm (Denos 1991) which automatically determines seed points in the top layer of an image pyramid, these are then matched using the Gruen adaptive least squares algorithms in the top layers of the pyramid whilst in the finer layers the Otto-Chau region matching algorithm, developed from the Gruen approach, is used.

The final stage is the transformation of disparities to 3-D data. In the current work the method used is that of Clark (1991) , which uses the range and Doppler equations for the two images to determine the vector P which is the co-ordinate vector of the ground point, using known satellite position S_1 .and S_2, and the satellite velocity. The principle is shown in Figure 1.

Clark has shown that the method is sensitive to errors in timing and that the method will be more stable with a longer base length. Work at UCL with ERS-1 data (Chen 1993) has shown that the method works but requires improvement to give satisfactory results.

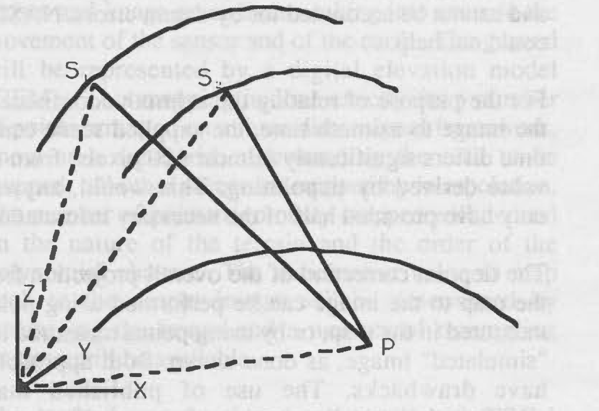


Figure 1. Principle of stereoscopic SAR.

5. SPECKLE REDUCTION FILTERS

For many applications of SAR data the presence of speckle causes severe problems. Because of the nature of stereomatching it was assumed that the same would be the case for that application, early tests by Denos (1991) with Seasat data showed this to be the case. Further ests were therefore carried out at UCL (Clochez 1993) to determine which filter, out of a number described in the literature, was most suited for this purpose.

Some initial tests were carried out on 12 filters in order to ensure that there was no serious distortion of radiometric values or degradation of edges. In order to test the effect of these filters on stereo matching the matching algorithm was applied to two overlapping sub scenes of PRI data covering an area in southern France to the north-west of Marseille. All of the scenes were fitered prior to matching. In order to evaluate the performance a number of criteria should be used:

- coverage - this is assessed by the number of successful matches;
- number of blunders;
- accuracy of stereo matching - this is assessed by the eigenvalue of the matrix produced by the least squares matching;
- distribution of the matches this can be assessed by a visual inspection of the matches;
- accuracy of resulting DEM - which can be evaluated by comparing the DEM which is derived from the disparities, with the DEM derived from other sources.

The coverage achieved for the filters is shown in table 1. These results were achieved with 5 tiers of the image pyramid, the use of a sixth tier slightly improved the coverage but a greater number of points were rejected. It can be seen that the best result is achieved without any filter but that there is little to choose between the filters as regards coverage.

FILTER	Rejected points	Coverage %
None	239	73.4
Lee	271	71.7
Kuan	248	72.6
Frost	241	67.7
Mod. Frost	247	69.1
α linear	255	71.0
MAP	264	70.7
β	250	72.4
γ	215	72.5
Li	251	70.8
LVN	258	71.0
Crimmins	264	68.2
log linear	250	70.8

Table 1. Coverage and rejected points with different filters

The blunders can be assessed by two methods, one is visual inspection and the other by maximum and minimum errors. Visual inspection shows that the number of blunders and the smoothness of the results varies considerably and that from this visual inspection the MAP filter appears to give the best result, however, detail is lost with this filter but the α linear filter retains the detail at the expense of more blunders. The worst results are, as expected, in the areas of highest relief.

The quality of match, as indicated by the eigenvalue also showed very little variation with different filters. Table 2 shows the maximum and minimum errors when the scaled disparity model is compared to the DEM of the test area. The standard error is very uniform but it can be seen tht the LVN, α linear, β and Lee produce the smallest maxima.

It can be seen then that the filters have the most effect on blunders but that the overall effect is small. It would seem that the best approach to further development is to improve the blunder detection routines within the stereomatching pyramid.

Table 2 also show that the accuracy produced after use of the various filters varies very little.

FILTER	Mean error	Standard error	Max. error
None	2.7	29.6	171
Lee	1.6	29.0	141
Kuan	2.1	30.4	176
Frost	1.1	29.7	202
Mod. Frost	0.8	29.7	201
α linear	2.2	29.7	135
MAP	0.9	29.2	162
β	1.9	29.0	143
γ	2.2	29.6	201
Li	2.9	30.0	166
LVN	1.2	28.6	122
Crimmins	1.5	29.3	169
log-linear	1.7	29.2	190

Table 2. Statistics relating DEM with disparity model.

6. RESULTS FROM DEM PRODUCTION

The final stage of DEM production is the transformation of the image disparities into three dimensional co-ordinates. This requires a geometric model as discused in section 2. Chen (1993) has developed the model described by Clark (1991) and has produced 3D co-ordinates without the use of ground control points from opposite side and same side stereoscopic pairs using roll-tilt mode data, of he area of Provence, France, to the North of Marseilles. The accuracy of the results has been tested with 38 ground check points. The method makes no use of ground control points but requires accurate orbit data. To date there have been difficulties in determining the accuracy of the orbit data and this may contribute to the errors in the results. Table 2 shows the statistics for the 38 points used to check the transformation.

It is apparent that there are systematic errors in both sets of data but that they apply to X and Z co-ordinates derived from same side data and only in Z with opposite side data. An analysis of errors indicates that these results are consistent with theory and that improved timing data and use of ground control points would give better results.

Leberl (1990) has reported results from stereoscopic SAR as shown in table 4. The current results indicate that the method with ERS-1 data needs improvement to give comparable accuracy.

Sensor	Easting (m)	Northing (m)	Ht (m)
Seasat	59	34	28
SIR A	48	84	93
SIR A	56	77	72
STAR*	13	26	28

*These results are from an airbourne system with GPS

Table 4. Results from stereo SAR systems quoted by Leberl (1990).

7. CONCLUSIONS

Geocoding of SAR is now carried out at a routine process. Experience has shown that processors developed for ERS-1 can also handle JERS SAR data but thet there can be problems in obtaining the correct timing and orbit data. These may be overcome with the use of tie points. It is important to include a system for validation into the geocoding process and this includes operator interaction. The production of auxilliaty products such as layover, shadow and energy maps is also important.

The work at UCL indicates that the use of stereoscopic SAR data can produce DEMs which can be used for geocoding. A SAR post proessing system which had provision for stereoscopic height determination and segmentation of the image into

	Same side pair			Opposite side pair		
	X(m)	Y(m)	Z(m)	X(m)	Y(m)	Z(m)
Minimum residual	-48	-91	-116	-85	-64	-103
Maximum residual	137	93	8	72	45	28
Mean	51	6	-42	4	-7	-51
Standard deviation	70	35	52	37	23	64

Table 3. Statistics from 38 ground check points.

relief areas for geocoding would give great flexibility for image fusion.

It appears that there is no clear advantage in any particular filter but that attention needs to be paid to blunder detection.

The sources of error in the stereo modelling process may be atrributed to a number of causes which include errors in azimuth timing and range. Such errors are not necessarily due to poor orbit determination or errors in the SAR processor but maybe due to the presentation and interpetation of the data in the header. Errors could also be due to weak geometry and measuring errors in the image. Future work will concentrate on making the optimum use of the header data available and improving the geometry of the viewing and tie point measurement. Roll-tilt data will be used and the model developed to include the provision for the use of more than two images. In the longer term other data such as SPOT could also be used in the same geometric model.

ACKNOWLEDGEMENTS

The data used in the work discussed in this paper was provided by ESA to University College London as principlal investigators for ERS-1 and as participants in the OEEPE/GeoSAR test of geocoding SAR data. The geocoding software was used with permission of GEC Marconi Research Centre who develped the software for the Defence Research Agency. The assistance of these three bodies, and of DLR, Oberpfaffenhofen is much appreciated.

REFERENCES

Chen Pu-Huai, 1993. Extraction of three dimensional data from European ERS-1 Synthetic Aperture radar imagery. MSc report, University of London. 82 pages.

Clark C, 1991. Geocoding and stereoscopy of Synthetic Aperture Radar. PhD Thesis, University of London.

Clochez O, 1993. Speckle noise reduction and stereomatching in ERS-1 SAR imagery. Training period report, UCL/ENSTBr Erasmus programme.

Denos M, 1991. An automated approach to stereo matching Seasat imagery. Proceedings of British Machine Vision Conference, Glasgow, 1991. Springer-Verlag

Dowman I, Clark C and Denos, 1992. Three dimensional data from SAR images. Int Arch. of Photogrammetry and Remote Sensing, 24(B4):425-427.

Dowman I, Laycock J, Whalley J, 1993a. Geocoding in the UK. SAR Geocoding:Data and Systems. Wichmann, pp 373-388.

Dowman I, Upton M, de Knecht J and Davison C, 1993b. Preliminary studies on the application of ERS-1 data to topographic mapping. Proc. of First ERS-1 Symposium, Cannes 1992. ESA SP-359:543-549

Dowman I, Chen Pu-Huai, Clochez O, Saundercock G, 1994. Heighting from stereoscopic ERS-1 Data. Proc of 2nd ERS-1 Symposium, Hamburg, 1993. ESA SP361,pp 609-614.

NASDA 1992, User's Guide for JERS-1 SAR Data Format, 1st edition.

Leberl F, 1990. Radargrammetric Image processing. Artech House Inc., Norwood, USA.

Renouard L and Perlant F, 1994. Geocoding SPOT products with ERS-1 Geometry. Proc of 2nd ERS-1 Symposium, Hamburg, 1993. ESA SP-361, pp653-658

Shimada, M, 1993, User's Guide to NASDA's SAR Products, EarthObservation Center, National Space Development Agency of Japan, HE93014 Rev.0..ex

Winter R, Kosmann D, Schulx B-S, Sties M, Wiggenhagen M, 1993. Radarmap of Germany - first mosaic and classification. Proc. of First ERS-1 Symposium, Cannes 1992. ESA SP-359:

Traitement de l'information transmise par les satellites d'observation des ressources de la Terre et le satellite japonais d'exploitation des ressources terrestres (ERS-1 SAR et JERS-1 SAR), équipés de radars de cartographie à ouverture synthétique

Résumé

Les données fournies par les radars à ouverture synthétique ERS-1 se sont révélées très stables et bien étalonnées et présentent un intérêt considérable pour plusieurs applications géoscientifiques. On a connu quelques problèmes avec la production des données JERS-1 mais celles-ci sont malgré tout très utiles. Le Collège universitaire de Londres (UCL) se sert des données recueillies par les deux capteurs pour le géocodage et pour la détermination des altitudes par stéréocorrespondance. La communication présente les caractéristiques des données provenant des deux capteurs ainsi que le traitement effectué pour le géocodage à l aide du logiciel mis au point à l'UCL et par un consortium dont faisait partie l'UCL. La dérivation des paramètres de géocodage dépend beaucoup du processeur employé pour la création des images ainsi que de la précision des données orbitales.

L'établissement de modèles numériques d'altitude à partir d'ERS-1 et de JERS-1 nécessite également des données orbitales mais dépend des caractéristiques radiométriques des données. La communication décrit le logiciel élaboré pour le filtrage et la stéréocorrespondance des données et présente les résultats obtenus. (Au moment de la rédaction, la stéréocorrespondance des données JERS n'avait pas été entreprise.)

La communication traite par ailleurs des problèmes rencontrés et des opérations à effectuer pour réaliser le géocodage et la stéréocorrespondance.

Évaluation des images RADAR pour la mise à jour de l'information géo-forestière

Alain Coulombe[1], Mario Hinse[2]

[1] Ministère des Forêts du Québec

[2] Ministère de l'Énergie et des Ressources du Québec

RÉSUMÉ

Dans le programme annuel du Service de la comptabilité forestière du ministère des Forêts du Québec, plusieurs centaines de milliers de dollars sont alloués à l'acquisition de photographies aériennes pour la gestion des inventaires forestiers. Afin de réduire les coûts entraînés par cette opération, des solutions alternatives sont explorées. Dans cet optique, l'utilisation d'images satellite de télédétection pourrait permettre des économies significatives au niveau de l'acquisition de l'information.

À l'heure actuelle, les images optiques (Landsat-TM ou SPOT) sont utilisées de façon opérationnelle pour certains aspects du traitement de l'information géo-forestière. Mais les images qu'on ne peut assurer aux gestionnaires du programme, qu'il y aura des bonnes images, à tous les ans et pour tout le territoire québécois. Ceci représente une contrainte importante, car les gestionnaires peuvent difficilement se permettre une absence aléatoire d'images dans le processus de mise à jour des inventaires forestiers.

Les récentes recherches dans le domaine des images RADAR suggèrent un potentiel d'application intéressant pour la mise à jour de l'information géo-forestière. Le RADAR a la propriété de pénétrer les couches nuageuses et il se soustrait aux mauvaises conditions météorologiques. On sait déjà que les images ROS aéroportées permettent une détection rapide des perturbations majeures (feux, coupes à blanc, coupes par bandes). Toutefois, plusieurs questions restent en suspend quant aux potentiels de détection des perturbations partielles du couvert forestier (coupe à diamètre limite, coupe de jardinage).

Les objectifs de cette étude sont (1) d'évaluer le potentiel des images ROS aéroportées (ROS C/X du CCT en bande C) pour la mise à jour annuelle de l'information géo-forestière, (2) évaluer le potentiel de l'hybridation des images ROS aéroportées avec les images optiques satellitaires en milieu forestier et (3) évaluer le potentiel de certaines images ROS satellitaires de types ERS-1 et J-ERS-1. Les images ROS aéroportées et les images de SPOT et Landsat-TM, utilisées dans le cadre de ce projet, sont des ortho-images résultant d'une correction géométrique rigoureuse développée au Centre canadien de télédétection.

Une première interprétation des images révèle que les scènes prises en hiver offrent une distinction plus franche des limites de coupes forestières comparativement à des images prises en l'absence de neige. Le couvert de neige aplanit et masque la régénération du parterre des coupes, permettant ainsi meilleure discrimination de ces dernières par rapport au couvert forestier avoisinant.

Evaluation of Radar Images for Updating Geo-Forestry Information

Abstract

As part of the annual program of the forestry accounting service of the Department of Forestry of Quebec, several hundred thousand dollars are allocated to the acquisition of air photos for the management of forest inventories. With a view to reducing the costs of this operation, alternative solutions are being explored. To this end, the use of remotely sensed satellite images could lead to significant savings in data acquisition.

Currently, optical images (Landsat-TM or SPOT) are used in operations for certain aspects in the processing of geo-forestry information. However, optical images are subject to the weather conditions prevailing from year to year, such that the managers of the program cannot be assured of good images every year for all parts of Quebec. This is a major limitation, as the managers can hardly tolerate random absences of images in the process of updating forest inventories.

Recent research in the field of RADAR images suggests a potentially advantageous application in the updating of geo-forestry information. RADAR is able to penetrate cloud cover and is unaffected by poor weather conditions. We know that airborne SAR images allow rapid detection of major disturbances (burns, clear cuts and strip cuts). However, there are still a number of unanswered questions about the potential for detecting partial disturbances in the forest cover (such as diameter-limit cutting and selection cutting).

The objectives of this study are (1) to evaluate the potential of airborne SAR images (SAR C/X of CCRS in C-band) for the annual updating of geo-forestry information; (2) to evaluate the potential of hybridizing airborne SAR images with satellite optical images in the forest environment; and (3) to evaluate the potential of certain satellite SAR images of the ERS-1 and J-ERS-1 types. The airborne SAR images and SPOT and Landsat-TM images used in this project are ortho-images resulting from rigorous geometric correction developed at the Canada Centre for Remote Sensing.

A first interpretation of the images reveals that the scenes taken in winter provide a clearer distinction of forest cut boundaries than do images taken when snow is absent. The snow cover smooths and masks regeneration in cutovers, making them easier to distinguish from nearby forest cover.

INVESTIGATION OF A GIS FOREST BASE MAPPING PROCESS USING OPTICAL AND RADAR REMOTE SENSING: FIRST RESULTS

Brian Davis
MacMillan Bloedel Ltd.
Woodlands Services Division
65 Front Street
Nanaimo, B.C.
V9R 5H9

Stéphane Rossignol
Radarsat International Inc.
3851 Shell Road, Suite 200
Richmond, B.C.
V6X 2W2

Christine Hutton
Canada Centre for Remote Sensing
Natural Resources, Canada
588 Booth Street
Ottawa, Ontario
K1A 0Y7

ABSTRACT

A technique was developed to integrate SPOT and simulated RADARSAT Fine Resolution Synthetic Aperture Radar (SAR) data into the MacMillan Bloedel Ltd. forest base map updating process. The project test site was located in the Queen Charlotte archipelago. Remote sensing data were corrected for terrain distortions and served to map new roads and cutblocks. The new features were added a GIS forest map based on the British Columbia 1:20,000 scale TRIM data. Much of MacMillan Bloedel's tenure in coastal British Columbia is constantly hampered by cloud cover which prevents the acquisition of optical satellite data. RADARSAT imagery with its all weather capability is expected to be an important source of information for base map updating operations. Methodologies developed support the concept of a future operational forest map updating system using SPOT data where available, and RADARSAT SAR data in areas where optical data is non-existent.

KEY WORDS: Remote Sensing, Synthetic Aperture Radar, SPOT, RADARSAT, ERS-1, DEM, GIS, Forest Base Mapping

1. INTRODUCTION

The Geomatics Team of Woodlands Services Division is responsible for the development and operation of MacMillan Bloedel Limited's Geographic Information System (GIS) including maintenance of the company's largest mapbase, the forest inventory. The ARC/INFO GIS is used to manage the 1:20,000 mapbase covering a combined area of approximately 1.4 million hectares. The mapbase was created by manually digitizing hardcopy planimetric maps that are inaccurate by today's standards. The original mapping sufficed for forest inventory purposes at the time but is no longer adequate due to the more stringent requirements of a GIS and associated applications.

MacMillan Bloedel intends to replace the mapbase with digital, three-dimensional topographic data based on the North American Datum 83 (NAD83) coordinate system. The British Columbia government's TRIM mapping program (Province of British

Columbia, 1990) is acceptable as a primary source of that data and, with the addition of features specific to MacMillan Bloedel's needs (cadastre, forest inventory), will form a new spatial database for the GIS. Hardware and software were acquired that will facilitate loading, enhancing and revising of the TRIM data while maintaining accuracy and topological integrity.

2. FOREST BASE MAPPING

2.1 Traditional Forest Mapping Procedure

Annual revision of the mapbase is an important function as MacMillan Bloedel harvests approximately 10,000 hectares of forest each year and the resultant land changes (new roads and cutblocks) must be added. Although this process is a corporate requirement, the responsibility for documenting these changes rests with each of the ten individual woodlands divisions. These divisions have traditionally used time-consuming field sketching techniques to update paper copies of the current mapbase that may or may not reflect the previous revision period(s). There has not been any standardized compilation techniques or understanding of data quality at this level. In addition, the logistics of coordinating the data sources, acquiring the data and creating source documents for digitization has been complicated and labourious. This method for performing the annual Inventory Revision has been identified as inadequate for the current mapbase and is totally unsuitable for the new spatial database. Aerial photography has never been implemented as an annual data source due to the cost and difficulties in coordinating flights over such a large area.

2.2 Forest Mapping and Satellite Remote Sensing

Satellite imagery, both optical and radar, has been identified as a viable data source for MacMillan Bloedel's map revision for a number of reasons. Firstly, the precise correction of satellite data to acceptable accuracies for operational mapping has become possible due to the availability of the digital TRIM data, especially the Digital Elevation Model (DEM). Secondly, the proposed revision process, by the very nature of the data source and its unique handling requirements, will be a centralized process with standardized compilation techniques performed by personnel with an appreciation for data quality and integrity. Thirdly, spatial coverage of the imagery will allow updating of large areas in one session at a considerably lower cost than with aerial photography. Fourthly, the re-visit capabilities of the satellites will permit revision on a more timely basis. Lastly, radar imagery in particular will permit all-weather acquisitions. This is an important factor as many of MacMillan Bloedel's operating areas are affected by the cloud and fog conditions that are typical of British Columbia's west coast climate.

The methodologies developed will allow MacMillan Bloedel to revise it's new spatial database in an economic and timely manner with an accuracy that, while not necessarily to TRIM mapping specifications, will be a known quantity that can be documented and applied to the decision making process. As satellite data spatial resolution improves, data compiled at less than the ideal specification can be replaced.

3. DATA SETS

Remote sensing data were collected over Louise Island in the Queen Charlotte Archipelago (52°56'N, 131°45'W). The data set consists of satellite Synthetic Aperture Radar (SAR), airborne SAR, simulated satellite SAR (using airborne SAR as input data), and satellite optical data.

The optical data set includes four bands of imagery from the French SPOT satellite. SPOT satellites carry two high resolution visible imaging systems (HRV). Each HRV system has two imaging modes: 1) the 10 m spatial resolution panchromatic mode (PLA) with a spectral range of 0.51 to 0.73 μm; 2) and the multi-spectral mode (MLA) operating in 3 bands at 20 m spatial resolution and spectral ranges of 0.50 to 0.59 μm, 0.61 to 0.68 μm and 0.79 to 0.89 μm. For the present study, PLA data was acquired May 27, 1993 and MLA data on August 21, 1992.

The SAR data set includes three flight lines from the CCRS airborne C/X SAR, one satellite SAR image from the European ERS-1 system, and two simulated RADARSAT Fine Resolution Mode images.
Narrow mode airborne C/X SAR data were acquired on July 26, 1993, in C-band (5.66 cm wavelength) and X-band (3.24 cm wavelength), dual polarization (HH and VV), with a nominal resolution of 6 m x 6 m and incidence angles ranging from 35° in near range to 76° in far range. Two of the flight lines were planned to correspond with RADARSAT orbital characteristics and look directions.

One ERS-1 SAR image was collected April 10, 1993, during the two week Roll-Tilt Mode (RTM) acquisition period. The ERS-1 SAR operates exclusively in C-band, VV polarization. The Roll-Tilt orbit phase permitted data collection with an incidence angle of 35° (mid-swath) and a nominal resolution of 30 m (The ERS-1 standard SAR mode operates at 23° incidence angle). The imaging geometry of the RTM data set is close to RADARSAT's Fine Resolution Mode (Figure 1).

RADARSAT Fine Resolution Mode simulation images were produced with C/X SAR HH polarized images as input data. The airborne data were modified using the INTERA SARPAC software package (Parashar and Wessels, 1989). These data were resampled to 9 x 11 m resolution and radiometrically altered to simulate the expected radiometric characteristics of RADARSAT imagery.

The base map information used is the BC Provincial TRIM data set. It consists of a digital elevation model (DEM) and digital planimetric information at 1:20,000 scale in the NAD83 coordinate system.

Figure (3a) shows the ortho-corrected SPOT PLA, (3b) one flight line of the airborne C-HH C/X SAR data (registered to the SPOT PLA), (3c) the TRIM DEM, (3d) and a composite image of the SPOT PLA (red), the C-HH C/X SAR (green), and the TRIM DEM (blue).

4. METHODOLOGY

The project developed separately for the optical and radar segments. The refinement of methodologies for the use of SPOT data was considered as the operational demonstration segment. The testing and development of techniques for the use of satellite SAR data was designated as the applied research segment.

Spatial resolution was the main criteria for data selection. The SPOT PLA band and the RADARSAT Fine Resolution Mode (simulation) were selected due to their 10 m and 11-9 x 9 m resolution respectively. Given new software developments for satellite ortho-rectification, it was felt that these sensors could reach acceptable mapping accuracies.

4.1 The SPOT Segment

The technique employed with the SPOT PLA data included:

A) Ortho-rectification and enhancement of the contrast between forest stands and cutblocks;
B) Integration of the ortho-corrected satellite data files with the GIS environment;
C) Interpretation of SPOT data for the extraction and mapping of new roads and cutblocks within the GIS environment.

4.2 The SAR Segment

Two main questions were investigated: 1) Can SAR effectively distinguish between forest stands, cutblocks and roads? 2) Can SAR meet the accuracies required for mapping these features in mountainous terrain?

The first question is being investigated using the airborne SAR and the simulated RADARSAT Fine Resolution data. The preliminary investigation of the airborne SAR has shown that cutblocks and new roads are distinguishable by their texture from the forest stands.

Investigations of the effect of radar incidence angle over forested sites have suggested that higher angles in orbital SAR imagery are preferred over steep angles to accentuate land cover information (Ahern and Raney, 1993; Rossignol and Ahern, 1992). The SAR data used in the present study will serve to further examine the influence of incidence angle for SAR forestry applications.

During the initial stages of this project, it was planned that RADARSAT mapping capabilities and accuracies would be tested with simulated RADARSAT Fine Resolution Mode data. After some reflection, the authors determined that absolute mapping accuracies of RADARSAT data could not be reliably estimated with simulated data. The imaging geometry of the input data for the simulations (airborne SAR data) is significantly different than what will be available with RADARSAT. Therefore, the airborne SAR data cannot effectively depict RADARSAT's geometric characteristics. Hence, the simulated RADARSAT data analysis was limited to a qualitative assessment of the sensor's ability to detect forest harvesting features.

Figure 1 and Figure 2 show the imaging geometries of the RADARSAT Fine Resolution Mode and the CCRS C/X SAR Narrow Mode data. The lower altitude of the airborne C/X SAR sensor produces a significant variation of incidence angle from the near to the far range within a single image (35° to 76° in our case). Near to far range incidence angle variation in orbiting SAR's are more subtle.

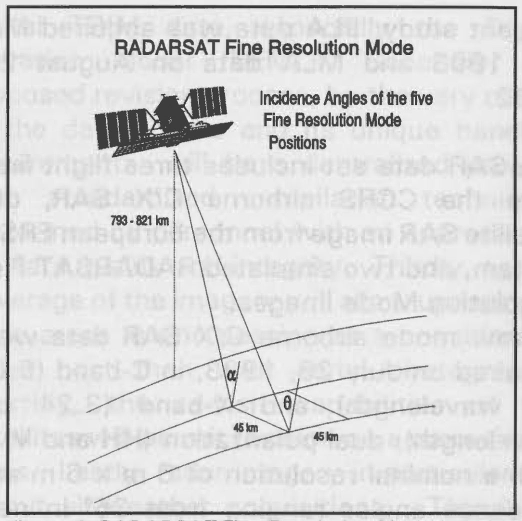


Figure 1 RADARSAT Fine Resolution Mode Imaging Geometry and Incidence Angle Positions

	α Near Edge	θ Far Edge
F1	36.9°	40.1°
F2	39.3°	42.3°
F3	41.6°	44.2°
F4	43.5°	46.0°
F5	45.3°	47.8°

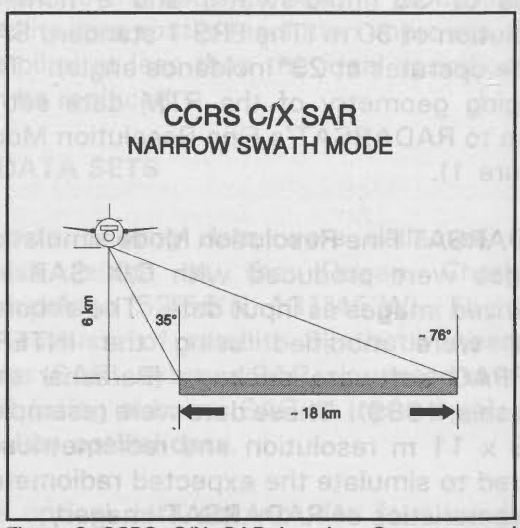


Figure 2 CCRS C/X SAR Imaging Geometry and Incidence Angles

The ortho-rectification process is being attempted with the ERS-1 Roll Tilt Mode image to develop an operational method for orbital SAR DEM correction. However, the accuracies of the ortho-rectified ERS-1 data are not considered to be fully representative of RADARSAT's attainable accuracies due to the inherent differences in resolution and imaging geometry between the two sensors.

Further analysis of the SAR data will include: 1) the merging of two flight lines with opposite look directions to eliminate radar shadowing in mountainous terrain; 2) radiometric rectification to correct for local incidence angle effects; 3) texture analysis to emphasize variations between land cover classes.

5. DISCUSSION AND RESULTS

The preliminary focus was on the SPOT segment of the project since at the time of printing, the SAR segment was at a preliminary stage. Results of the SAR investigation will be presented in the symposium session and will be reported in a later publication.

The SPOT ortho-rectification results are described in some detail as the availability of the ortho-rectification software and digital elevation information were unique in the authors' experience.

The ortho-correction procedure recently developed at the Canada Centre for Remote Sensing (CCRS) (Toutin and Carbonneau, 1992) and transferred to PCI software (PCI, 1993) for commercialization was employed. Simply described, the technique uses a four step process: 1) reading the digital & ephemeral data (satellite orbital and path information) from tape; 2) collecting ground control points (including elevation information); 3) calculating the algorithm; 4) and correcting the data.

The ortho-rectification package is compatible with a number of the processing levels provided by the satellite data vendors (RADARSAT International Inc. in Canada). In the first attempt at correcting the SPOT PLA and MLA data, georeferenced data was used. Once ortho-corrected, these data yielded Root Mean Squared errors (RMS errors) of 22 m in the Eastings direction and 6 m in Northings.

The MLA data set were included in the research effort to allow MacMillan Bloedel staff to familiarize themselves with digital classification and land cover discrimination techniques. The RMS errors achieved with the MLA data were 17 m in Eastings and 11 m in Northings.

An effort to improve the correction was sought. Discussions with the author of the correction algorithms suggested improvements to the accuracies could be expected the input data was restricted to the raw or bulk processed levels.

The authors are confident that a second attempt at ortho-correction with the more appropriate input data will increase the accuracies.

6. SUMMARY - PLANNED FUTURE WORK

The use of SPOT PLA for detecting and accurately mapping forest harvesting features has been demonstrated. Ortho-correction of the SPOT data, using the raw and bulk products as input, will be attempted. Current results are encouraging and are expected to fulfil MacMillan Bloedel's stringent requirements.

Work will continue to further improve the correction and information extraction techniques. SPOT and SAR mapping results will be compared with traditional photogrammetric techniques.

SAR data analysis will include the development of radiometric and geometric correction methodologies in addition to the investigation of texture analysis algorithms to accentuate variations between land cover classes.

The radiometric and geometric correction of the simulated RADARSAT and the ERS-1 data will serve as a demonstration for the possible capabilities of RADARSAT data. Although the results of the mapping accuracies with SAR data are not directly applicable to the RADARSAT case, the exercise should demonstrate the general capabilities of radar inventory mapping. Given the knowledge and experience acquired during this project, the Woodlands Services Division of MacMillan Bloedel will be well prepared to incorporate RADARSAT in their mapping process when the data become available.

ACKNOWLEDGEMENTS

C/X SAR airborne data were provided by CCRS. SPOT and ERS-1 data were provided by RSI. The authors also wish to thank Dr. Frank Ahern of the CCRS Forestry Applications Development Section for his comments and suggestions during the initial stages of the study and to thank Dennis Nazarenko of RSI for reviewing this document.

REFERENCES

Ahern, F.J., R.K. Raney, 1993, "An Almaz/ERS-1 Comparison Demonstrates Incidence Angle Effects in Orbital SAR Imagery", Canadian Journal of Remote Sensing, Vol 19, No. 3, pp. 259-262.

Boulton, S., B.G. Davis, S.C. Lanoix, R.A. Newsome, L.N. Rumming, and B.T. Whitehead, 1992, An Evaluation of MacMillan Bloedel Limited's GIS Map Base.

Davis, B., S.C. Lanoix, and L.N. Rumming, 1992, GIS Task Force Subcommittee Recommendations on Hardware and Software Purchases.

Livingston, C.E., A.L. Gray, R.K. Hawkins, R.B. Olsen, 1988, "CCRS C/X-Airborne Synthetic Aperture Radar: an R and D Tool for the ERS-1 Time Frame" in: Proceedings of the IEEE National Radar Conference, April 20-21, The University of Michigan, Ann Arbor, Michigan, pp. 15-21.

Parashar, S., G.J. Wessels, 1989, "Synthetic Aperture Radar Simulation Software Package at CCRS", Proceedings of IGARSS '89 / the 12th Canadian Symposium on Remote Sensing, Vol.3, Vancouver, pp. 1729-1732.

PCI, 1993, PACE Satellite Image Ortho and DEM, Version 5.2, Richmond Hill, Ontario.

Surveys and Mapping Branch, Ministry of Crown Lands, Province of British Columbia, 1990, British Columbia Specifications and Guidelines for Geomatics, Digital Baseline Mapping at 1:20,000, Release 1.0.

Toutin, T., Y. Carbonneau, 1992, "La création d'ortho-images avec MNE: description d'un nouveau système", Journal Canadien de Télédétection, Vol.18, no.3, pp. 136-141.

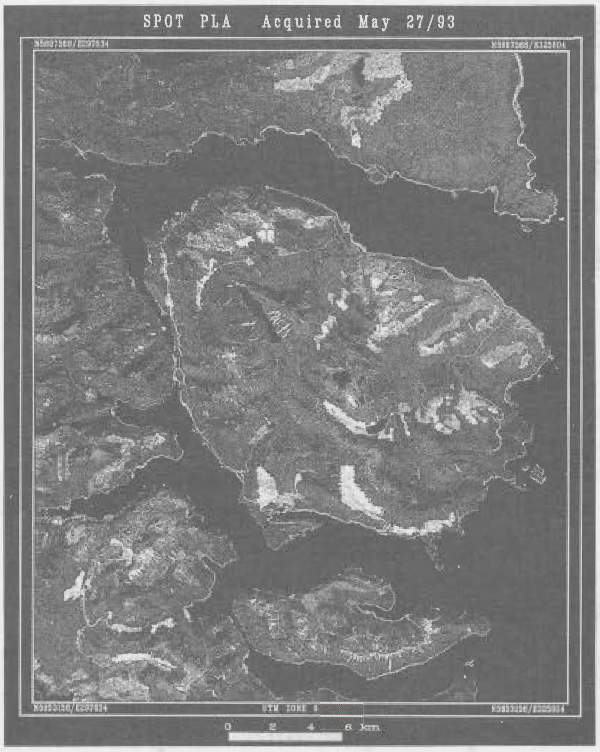


Figure 3a SPOT PLA ortho-rectified image

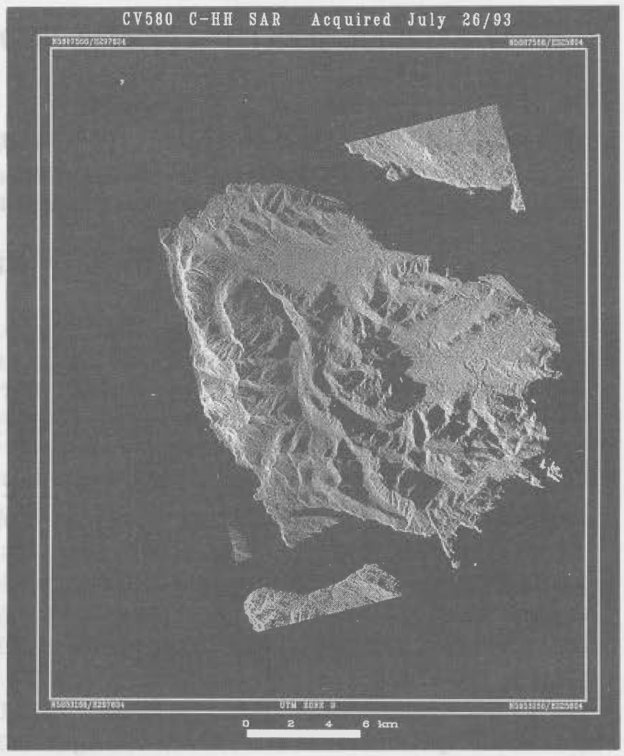


Figure 3b CCRS C/X C-band HH SAR, registered to the SPOT PLA image

Figure 3c B.C. Trim Digital Elevation Model

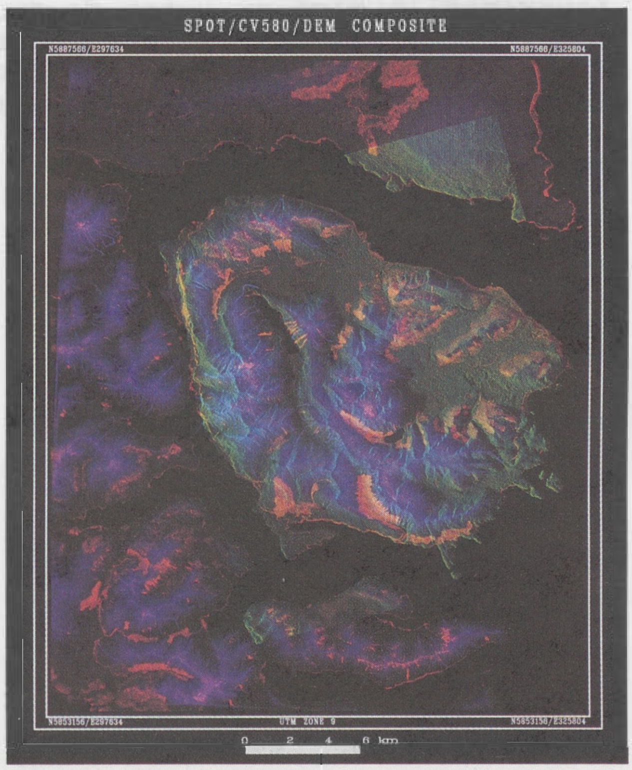


Figure 3d SPOT / C/X SAR / TRIM DEM Composite image; The SAR image (in green) does not perfectly overlay with the other data sets. For high relief areas, ortho-correction is required to correct for inherent SAR terrain distortions.

Étude d'un processus de mise à jour de cartes forestières SIG au moyen de données de télédétection optique et radar : Premiers résultats

RÉSUMÉ

On a mis au point une technique d'intégration des données SPOT, ainsi que des données simulées du radar à synthèse d'ouverture (R.S.O.) du satellite RADARSAT à un système de mise à jour des fonds de cartes des régions forestières de la MacMillan Bloedel Ltd. L'emplacement choisi en vue de ce projet était situé dans l'archipel de la Reine-Charlotte. Les données de télédétection ont été corrigées pour tenir compte de la distorsion engendrée par les formes de terrain et ont servi à cartographier les nouvelles routes et les zones récemment déboisées. Cette nouvelle technique a permis de mettre à jour les cartes d'inventaire forestier, mises au point grâce au programme TRIM, à l'échelle de 1/20 000. La plus grande partie des terres que détient la MacMillan Bloedel dans la zone côtière de Colombie-Britannique est constamment cachée par une couche nuageuse qui rend pratiquement inopérants les systèmes de télédétection optique par satellite. Cependant, l'imagerie RADARSAT, exploitable par tous les temps, devrait fournir une importante source de renseignements supplémentaires susceptibles de faciliter la mise à jour des cartes. Les méthodologies mises au point dans le cadre de ce projet devraient favoriser la mise au point de nouveaux systèmes de mise à jour des fonds de cartes des régions forestières à l'aide des données SPOT, dans la mesure où celles-ci sont accessibles, et aussi grâce aux données de télédétection par radar à synthèse d'ouverture du RADARSAT, pour les zones qui ne se prêtent pas à la télédétection optique.

MATHEMATICAL PROBLEMS OF REAL–TIME MAPPING AND DATA BASE MODELLING

Erhard Pross
Institute of Applied Geodesy
Leipzig, Germany

Commission II, Working Group 1

Abstract

In real–time processing the time is the main processing direction controlled by discrete operators. Functional analytical methods integrating discrete and continuous description and methods lead to constructive tools in data base modelling and real–time mapping. The realization of image operators as difference operators depends on the the fast convergence and on the techniques' stability and algorithmic ideas of switching between spatial and time iterations and processing directions.

KEYWORDS: database modelling, theory, scene analysis, optical flow

1 Introduction

With the triumphal progress in computer techniques in several fields also a digital photogrammetry and real–time mapping was formed. Numerous ideas and principles in physics, electronic engineering, and new disciplines like robot and machine vision influence the digital photogrammetry in a very great amount. So we must recognize not only an integration of geodesy, photogrammetry, and cartography in geoinformatics but also new concepts and models based on modern mathematics. The separate developments in modelling must be integrated in one model (see [1]). The integration of different sensor types in one real–time mapping model from the data capture to the the storage and the post–processing of the data is investigated in [4].

The idea of this paper is to view modelling in real–time mapping from an abstract mathematical point of view and give some directions to the algorithmic realization of techniques.

2 Mathematical Background of Data Base Modelling

2.1 Abstract Space Definition

For the definition of an abstract mathematical space corresponding to tasks in geoscience it is necessary to define measurements of nature signals as elements of such a space and technological processes and technical course as corresponding transformations between suitable spaces. Spaces are characterized by the contents (images, features, ...), by the discreteness (analog, digital) or by the structure (vector, raster).

The art of abstract mathematical modelling leads to such models, the concrete realization of which leads to special views on projections into and integrations within function spaces. The functional analysis is a mathematical building with powerful tools to realize such a modelling. In geoinformatics we know image spaces — including images, image sequences, raster maps, ... and feature spaces (see [3], [10]).

A mathematical model demonstrating a functional analytical description in geoscience especially remote sensing, photogrammetry, and cartography in one calculus was developed (see [7], [8]).

From the semantic level point of view image spaces have a low level and the feature space a middle or high level.

Coordinates are the spatial x–, y– and z–coordinate — in topography especially in smallscale topography the third coordinate (height) plays an important role by 2.5–, 2+1– or 3–D–GIS–modelling —, the features or attributes m and last but not least the time t.

The significance of the time leads to qualitative different models in regard of time

- low influence — time as date or up–dating number
- common coordinate
- dominant processing parameter for instance in real–time mapping

These several levels differ with different methods and techniques. For that reason the common methods in photogrammetry, remote sensing, and GIS — corresponding to the first and second level — cannot be transferred to real–time mapping processing.

A functional is a transformation of an abstract space

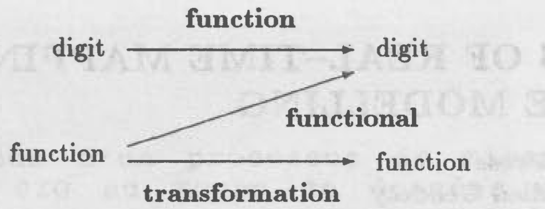


Figure 1: Function, Functional and Transformation

into complex numbers, i.e. if the space is a function space every function corresponds to a digit and the functional analysis is the analysis in these spaces (see also figure 1).

The functional analysis includes a topological and an algebraical part. The topology is realized by neighbourhood definition and measures. On the other hand with algebraical means we can define basic operators like addition, scaling, and multiplication. This is very important for the realization and also for modelling of data bases and real–time mapping.

The standard measure in mathematics and physics are normes $\| . \|$ and scalar products $\langle . , . \rangle$. The HILBERT–space methods — using the scalar product as a special norm — are powerful and constructive tools in algorithmization. The transformations within HILBERT spaces can be designed and analyzed by scalar products realized by integrals or series in the sense of L^2– and l^2–HILBERT spaces. Than the image space transformations (T_{In} — integration, T_{Pr} — projection) are modeled:

$$T_{In} = \langle . , \varrho \rangle_A$$
$$T_{Pr} = \langle . , \delta \rangle \quad .$$

ϱ is a weight-function corresponding to an area A and δ is the DIRAC distribution depending on the projection direction (see also [9]).

Numerous examples in image processing demonstrate this fact (duality between spatial and frequence spaces especially discreteness and periodicity, using the orthogonality by decoupling of coordinate–directions in integral transformations). Completed by algebraic principles several approximative techniques were developed.

The discretization in spatial coordinates and in time is the basis for computer realizations. This digital modelling leads consequently to the digital scene analysis. A great amount of processes in nature are described by differential equations and their digital analogon — the difference equations. Typical mathematical methods in real–time mapping processing are fast converging iterations of difference operator equations. In [2] this is demonstrated by a comparison of variation with least squares methods.

2.2 Scene Analysis Modelling

For a photogrammetric modelling in scene analysis it is necessary to generalize the description in time direction.

There are three time levels (point–series–continuity) with a qualitative jump from the discreteness to the continuity connected with new objects like image cube and trajectories. This includes also a new mathematical description level characterized by variation methods and functional analysis. In figure 3 this fact is shown in a symbolic kind.

It is typical that the search for a solution of problems in science and technology corresponds to finding the extrema of functionals. Often such functionals are measures of the difference of structures (e.g. least square means and variation methods). In [6] such a calculus in computer vision theory is demonstrated.

3 Problems of Real–Time Processing

The task of scene correspondence is to define the geometric relations between two or more images of the same domain. Thereby the time is running continuously and typically the time points are discrete on the time axis in the image space. Consequently these images also correspond to discrete levels in the image space (see also figure 3).

The correspondence problem is solved with the help of image information (grey values, image frequencies, textural features etc.) by using the variation calculus to minimize corresponding functionals. The minimizing problem leads to the solution of EULER equations, e.g. systems of partial differential equations (see also [2]).

Because a sequence of images is a set of discrete planes in the image space one direction of generalization is the change over to the time continuum and to get a complete image space. In this model it is possible to define termini as image flow or trajectories of objects. The transformations act nearly exclusive in the image space. As results of the discretization of the image space we get sequences of points. The correspondence is defined by the solution of attached difference equations being also an analogon in the digital sense to the differential equations. By using an iterative scheme it can be formulated as a complete digital photogrammetric task (see [2]). Further image processes can be designed by finding attached transformations within the image space especially by separation and qualified algorithmization in geometry and time. In dependence on the complexity of algorithmization the use of principles of parallelization is possible (see [5]).

Figure 2 shows the difficulties at working with image flows and digital images in one calculus because from

time \ space	continuous	discrete
continuous	image flow	digital image flow
discrete	image sequences	digital images

Figure 2: Structure Characterization

the point of view of structure categories the discreteness and the continuum are different.

Only an adaptive problem–depended digitalization in space and time leads to a regular grid structure for difference equation in x_i, y_j and t_k:

$$\Delta(x_i^{(l)}, y_j^{(m)}, t_k^{(n)}) = 0$$

By separation and with a multiindex (s) this equation is transfered into

$$x^{(s)} = \tilde{\Delta}(x^{(s-1)}) \quad .$$

In an example in [2] comparing variation methods with least squares methods this technique is demonstrated.

4 Outlook

The modelling in real–time mapping encloses the definition of suitable photogrammetric principles of conservation (image intensity, image energy, velocity field, ...) and corresponding realizations, which are as a rule linear combinations of mathematical operators like an equation of weighted sums of operators, that must be effectively solved by iteration of difference equations determined by image convolutions.
By using several images or a quasi–continuous image cube methods in real–time techniques essentially differ. Another view to modelling realizes the image cube trajectories as parallaxes. The velocity is the first derivative at the time and the velocity variation is the second derivative or the acceleration. The parallax can be interpreted as result of a projection in time direction within the image cube — also the digitized image flow corresponds to the parallax — and this results in the discretization distance in time direction being rougher and also influencees the switching strategy between processing direction in an algorithmic solution.

I hope that this paper could contribute to the discussion of mathematical principles , methods, and background in modern mapping techniques and data base modelling.

References

[1] Foerstner,W.: Objektmodell fuer Fernerkundung und GIS, Jahrestagung der DGPF, Cologne, 1991

[2] Hahn,M.; Pross,E.: Bildzuordnung nach dem Variationsprinzip – ein Vergleich mit Kleinste–Quadrate–Methoden, Zeitschrift fuer Photogrammetrie und Fernerkundung 4(60), pp. 116–123, Karlsruhe, 1992

[3] Leberl,F.: The Promise of Softcopy Photogrammetry. In: Ebner, Fritsch, Heipke (Eds.): Digital Photogrammetric Systems, Wichmann, Karlsruhe 1991, pp. 3–14

[4] Novak,K.: Real–Time Mapping Technology, Intern. Archives of Photogrammetry and Remote Sensing, Vol. 29-2, Washington 1992, pp. 569–575

[5] Mueller,B.; Hahn,M.: Parallel Processing — The Example of Automatic Relative Orientation, Intern. Archives of Photogrammetry and Remote Sensing, Vol. 29-2, Washington 1992, pp. 623–630

[6] Poggio,T.; Torre,V.; Koch,Ch. (1985): Computational Vision and Regularization Theory, Nature 317, pp. 314–319

[7] Pross,E.: Describing Digital Photogrammetry by Functional Analytical Methods, Intern. Archives of Photogrammetry and Remote Sensing, Vol. 28-2, Dresden 1990, pp. 336–344

[8] Pross,E.: Digitale Komponenten von Fernerkundung, Photogrammetrie und Kartographie innerhalb der Geoinformatik sowie deren Beschreibung durch funktionalanalytische Mittel und Methoden, Habilitation, Dresden 1991

[9] Pross,E.: Die Bedeutung der Funktionalanalysis bei der Integration von Geowissenschaften fuer Umweltaufgaben, Veroeff. Zentralinstitut fuer Physik der Erde. No. 118, Potsdam 1991, pp. 353–362

[10] Pross,E.: Importance of Modern Mathematical Methods for the Digital Photogrammetry, Intern. Archives of Photogrammetry and Remote Sensing, Vol. 29-2, Washington 1992, pp. 369–373

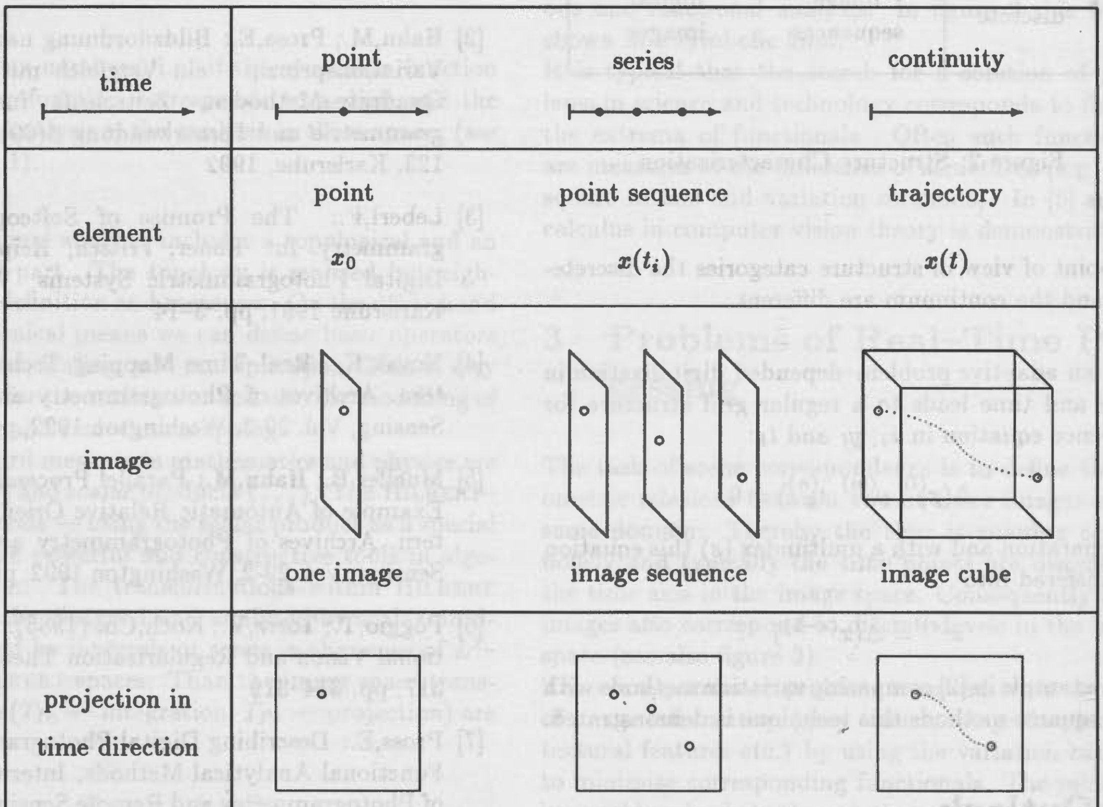


Figure 3: Time Characterization

Problèmes mathématiques de cartographie et de modélisation de base de données en temps réel

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif.

IDENTIFICATION AND LOCATION OF SIMPLE OBJECTS FOR REAL-TIME MAPPING

Christoph Geiselmann
RIB Bausoftware GmbH
Vaihinger Str. 151
70567 Stuttgart / Germany

Michael Hahn
Institute of Photogrammetry
Stuttgart University
P.O.B. 106037
70049 Stuttgart / Germany

ISPRS Commission II , Invited paper

KEY WORDS: Classification, real-time mapping, invariant features, affine-invariance, colour image sequence, motion based segmentation, traffic signs

ABSTRACT

Systems for real-time mapping can be described like most information systems by components for recording, storage, management, analysis and presentation of the data. The recording and storage of data collected by positioning and imaging sensors is done in real-time. Up-to-date real-time is not fully achieved by the other system components. In particular the analysis of the images is often complex and time consuming.

In this paper, an investigation in the identification and location of simple objects is presented. As an example traffic signs recorded by a colour image sequence are used. As a by-product this study shows a possibility on how to reduce the user activity considerably in the process of creating traffic sign maps. The developed automatic procedure consists of the following steps: a motion stereo pair is used for image segmentation into a region of interest and in background. As a result a precise displacement vector field of this region is derived. The deconvolution of motion blur leads to a restoration of the image and prepares the image for further processing. Next the borderline of the region is determined and affine-invariant features of this contour are determined. For interpretation the suitability and separability of the features is of special interest. Finally, maximum likelihood classification with these features leads to the identification of object classes for the localized traffic signs. The reported experiments in this investigation are based on simulations and on real world scenes.

KURZFASSUNG

Systeme zur Echtzeitkartierung lassen sich wie Informationssysteme generell durch Komponenten charakterisieren, die die Aufnahme, Speicherung, Verwaltung, Analyse und Präsentation beinhalten. Die Aufnahme und Speicherung der Daten, die von Positions- und Bildsensoren erfaßt wurden, erfolgt in Echtzeit. Dies wird bislang von den anderen Systemkomponenten, insbesondere der Bildanalyse, noch nicht erreicht.

Dieser Aufsatz präsentiert eine Untersuchung zur Identifikation und Lokalisierung einfacher Objekte in Farbbildsequenzen, wofür Verkehrszeichen ausgewählt wurden. Dabei wird ein Weg aufgezeigt, wie das interaktive Messen durch den Anwender bei der Verkehrszeichenkartierung beträchtlich reduziert werden kann. Die entwickelte automatische Prozedur beinhaltet eine Bildsegmentierung über die Bewegung innerhalb aufeinanderfolgender Bilder. Dabei wird das Interessensgebiet vom Hintergrund getrennt und ein genaues Verschiebungsvektorfeld für dieses Gebiet bestimmt. Die Bewegungsunschärfe im Bild kann durch Bildrestauration eliminiert werden. In einem weiteren Schritt wird die Begrenzungslinie dieser Region extrahiert und durch affin-invariante Merkmale beschrieben. Die Eignung und Trennbarkeit dieser Merkmale für die Interpretation ist von besonderem Interesse. Mit diesen Merkmalen gelingt es, durch Maximum-Likelihood-Klassifizierung die Verkehrszeichen bezüglich ihrer Objektklassenzugehörigkeit zu identifizieren. Für die experimentellen Untersuchungen wurden Simulationen und Realweltszenen herangezogen.

1. INTRODUCTION

Today real-time mapping concentrates on two topics: the first topic addresses the combination of data collection sensors like positioning and imaging sensors for digital mapping. The second topic is concerned with the development of mapping procedures, i.e., methods and algorithms for the identification and location of objects which have to be captured and stored in a database. Because of the real-time aspect of mapping simple but reliable procedures are required. With regard to this second topic practice looks quite different. A mapping equipment enables the *user* to identify and locate objects in an absolute reference frame. The automatic identification of objects is still a hard problem of vision research.

Real-time classification procedures are most successful in very restricted industrial scenes. In this case the number of object classes is very small. By thresholding the images the objects are separated from the background. The extraction of some simple features often is sufficient for the identification of the objects. Significantly more demanding are outdoor scenes, where in general a lot of objects are present and the scene is complex.

This paper focuses on the identification of simple objects in outdoor scenes. As an example we use traffic signs and assume that the objects of interest are located in a more or less natural environment. Such a scene might be captured by a Highway Inventory System (Schwarz, 1992) or similar surveying vans (Novak, 1990). For the experiments in this paper we use a short sequence of colour images. The images are taken by a standard video camera, i.e., the full frames (25 Hz) are composed by the odd and even fields of the half frames. Thus a full frame is the smallest motion unit in which information about the scene is captured and, in addition, the effects of the movement of the car are represented.

The geometric structure or shape of the objects of interest is simple. They can be modelled by a plane which is spatially limited, for example, by triangular, rectangular or circular border lines. Traffic signs are typical objects of this class. For the recognition the colour of those objects can be expected to be an important clue. The use of a small number of different colours like red, yellow, blue, black and white keeps the discrimination between objects relatively easy.

In developing a procedure for the identification of objects from an image sequence some specific problems have to be taken into account. In general, the size of an object in the image is unknown. The projection of the object leads to a perspective distortion of the border line. Thirdly the movement of the video camera during the exposure is the reason for motion blur.

An important point for solving the mapping task by an automatic procedure is the efficient use of different information sources like the colour, motion and contour of the object. Simple algorithms are required to achieve real-time capabilities. The procedure we propose for the recognition of simple objects consists of the following steps. First regions of interest, i.e. regions which may contain the unknown objects, have to be detected. Although in principle colour, motion and contour may contribute to solve this task only motion is used in this step. The result of the motion segmentation is a displacement vector or displacement field of this region. The displacement field can be used to restore the image with the aim of eliminating the effect of motion blur. The restored image is well prepared for the extraction of the border line of the object by which the location of the object in the image is determined. Because of the 3D to 2D projection of the object its border line is distorted. In consequence, for recognition it is advantageous to extract affine-invariant features from this contour. The last step of the procedure is a maximum likelihood classification. The affine-invariant features can be used to determine the most likely object class for the unknown localized object. In the following these main steps of the analysis are discussed further. Concerning the interpretation the suitability and the separability of the object features is of special interest. Therefore, this aspect will be analysed in more detail.

2. DETERMINATION OF THE REGIONS OF INTEREST

A first step in the recognition of an object is to identify a region in the image in which the searched object is supposed to appear. This region is called the region of interest. In the case of a colour image sequence at least three different possibilities exist to detect a region of interest.

Objects like traffic signs are placed at distinguished places, for example, along the street with a certain distance from the background. The movement of the camera relative to the scene usually leads to a relatively large motion of the object in the image. This motion in the image space is one source for separating a region of interest from the background. Another source is the known colour of the object. Segmentation procedures based on colour benefit extremely from distinct colours which are unique in the real world scene. A third possibility is to use the structure of the object. The border line with its edges and corners is a very characteristic geometrical feature for those simple objects. On the whole these three possibilities are largely independent. Consequently, if all three sources are exploited in a cooperative manner a most reliable result can be expected.

In Geiselmann (1992) all three possibilities are investigated. The procedure using motion has been experimentally most successful. Therefore the segmentation based on motion is described in more detail in this section.

If the location of an object in the image changes then the intensities of the corresponding region will change in general too. Hence close at hand is to detect intensity changes between images and identify these pixels as candidate elements for the region of interest. Unfortunately also other factors like changes in the illumination and noise lead to changing image intensities which makes motion detection more difficult.

According to Jain (1981) two major types of approaches for motion detection can be distinguished. The simpler ones calculate difference images, detect regions of changing intensities and try to estimate the displacement, for example, by correlation. If the displacement is significant then motion is detected. More complex are procedures of the second group. A symbolic description with points, lines and regions is extracted in each image individually. Various proposals for establishing feature correspondences exist based on hierarchical processes, on prediction, on relaxation or on more simple procedures like displacement histograms.

Procedures relying on difference images are faster and in most cases have smaller storage requirements than the other ones. They more likely agree with the requirements of real-time mapping, thus we prefer a approach of this type. Our procedure for the extraction of regions of interest is sketched in figure 1. Prerequisites are a densely sampled image sequence and a minimum size of the object in the image of, for example, 100 pixels. Further is assumed that the object is not occluded. The single processes of the procedure (figure 1) deal with problems resulting from sensor noise and from small changes in the illumination within the natural environment. The object may overlap or not in the difference image. Smaller deviations

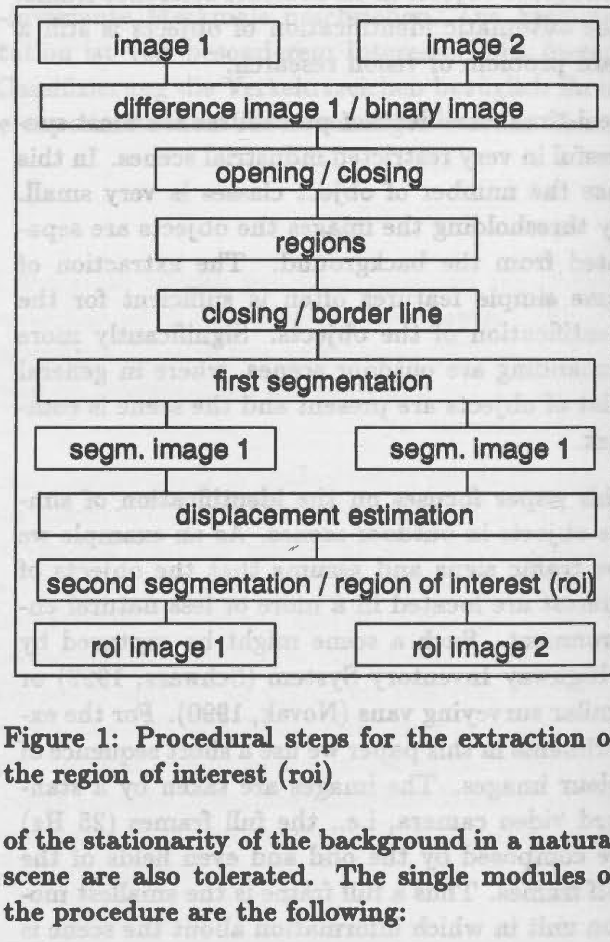


Figure 1: Procedural steps for the extraction of the region of interest (roi)

of the stationarity of the background in a natural scene are also tolerated. The single modules of the procedure are the following:

Difference image

Starting point are two subsequent images of the image sequence. In the experiments the both half frames of a video image are used. The intensity differences between both images are calculated pixel by pixel. Thresholding the intensity differences leads to the binary difference image. A value 1 in the binary image indicates a change, a value 0 tells that the intensity difference was smaller than the threshold. The threshold has to be found experimentally. Choosing the threshold is a weak point of the procedure.

Morphological operations

The binary difference image will represent at least one large region with the unknown object. In addition, noise and other effects will lead to a lot of small regions. These small regions are eliminated by 'opening', i.e. the binary image is processed by morphological operations. By 'closing' small gaps are filled and the contour is smoothed.

Regions

The explicit determination of regions in the binary image is done by a simple region growing, in which the 8-neighbourhood of the points is checked sequentially. If the opening has not removed all small regions this can be done in this step by thresholding the extracted regions with a required minimum area of a region.

Border line

A second closing with a larger operator size aims at filling gaps in the region of interest. Further filling is obtained by the following procedure. If background pixels appear within a horizontal line between object pixels they are turned to object pixels. The same algorithm is applied the vertical lines. The border line then will appear as a smooth closed contour of the region of interest. The border line is simply extracted by binary edge detection and edge linking.

First Segmentation

The border line circumscribes a region in the difference image which includes the object in both images simultaneously. Because of the displacement of the object during the acquisition of the sequence in each region of both images some background is included. In the first segmentation step the regions within the border line are extracted in each of both images separately.

Displacement estimation

The estimation of the two-dimensional displacement or, more generally, the displacement vector field between the segmented regions of both images can be done by cross-correlation or other well-known matching techniques. The search for the corresponding parts of both regions has to take into account that in both regions different parts of background are included.

Second Segmentation

The displacement vector is used to eliminate the background in both regions individually. This is achieved simply by extracting those parts in both regions which are identified as corresponding areas within the matching step. The result of this second segmentation step is the region of interest which is identified and located in both images individually and which only captures the object.

Within the whole task of object recognition the detection and location of the region of interest is an important process. An example which shows the single steps of the procedure is shown in figure 2. Just to give an idea on some experimental data the thresholds and other quantities are listed. For the first opening and closing a 3×3 operator is used. The threshold for the determination of the binary difference image was 16 grey values. A minimum size of an object of 50 pixels was required and in the second closing a 9×9 window was used. The estimate for the displacement was 1 pixel vertically and 8 pixels horizontally. The result of the processing is the region of interest (the two pictures in the lower right of figure 2). For the identification of the object this region has to be analysed further.

3. DETERMINATION OF INVARIANT FEATURES AND CLASSIFICATION

The extracted region of interest is the input for the identification the unknown object. As stated earlier an intermediate process is proposed to eliminate motion blur. Because motion blur can be avoided by using shuttered video cameras this step of processing is not obligatory. Thus a discussion of this topic is omitted in this paper. A description of the techniques for the deconvolution of motion blur can be found in Geiselmann (1992).

From a classification point of view the extracted regions are the candidates for the identification step. The task is to extract proper features from the candidates and to compare them with the features of the known objects. A result of the motion based segmentation of the image is the border line. This closed contour can be considered to be the observed geometric quantity of the unknown object. The problem is now that the appearance of the object in the image, and this applies to the border line as well, is distorted by the perspective projection of the object. If it is possible to find descriptive quantities which are not affected by the projection then object recognition is possible without knowing the rotation and translation between a camera centered and an object centered coordinate system. These quantities are called invariants (Forsyth et al., 1990).

Depending on the complexity of the object and on the formulation for deriving the invariants the theory of invariants defines exactly the number of invariant features with respect to a specific geometric transformation. For solving the recognition task by automatic pattern recognition often a few invariant features are sufficient. For the simple and planar geometric figures of the objects of interest suitable invariant features have to be found. In our application the object models for the traffic signs are restricted to triangles, rectangles and circles.

Usually, in photogrammetry the mapping from object space to image space is postulated to be a general perspective projection. The perspective projection can be approximated by an affine transformation for planar objects and if the distance between the object and the camera is large and the field of view is narrow (Costa et al., 1989). For an affine mapping it is easier to determine invariant features than for a perspective transformation. In the next section first some different procedures for the determination of affine-invariant quantities from points, lines and areas are discussed. For the simple objects only a few invariant features can be determined. Further the figures possess specific symmetries. Thus for these objects it is very interesting to find out which of the procedures are suitable for the determination of affine-invariant features at all. The result is a pre-selection of the possible procedures for the determination of affine-invariant features from a closed border line. Because a procedure should be as reliable as possible the sensitivity of the affine-invariant features with respect to noise in the border line is evaluated. A quality measure for the sensitivity is the probability for a wrong classification. This probability can be used for the assessment of the separability of parametric models based on the observations.

3.1 Determination of affine-invariant features

Various procedures are proposed for the determination of affine-invariant features using points, lines or areas. These features are affine-invariant coordinates (Costa et al., 1989) for the points, affine-invariant curvatures (Cyganski et al., 1987) or affine-invariant Fourier descriptors (Arbter, 1989) for the lines and affine-invariant moments (Hu, 1962) for the areas. The procedures will be shortly discussed in the following. The detailed mathematical formulations can be found in Geiselmann (1992) or in the references quoted above.

3.1.1 Affine-invariant coordinates

At least 4 points $(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})$ are necessary for the determination of affine-invariant coordinates. With 3 points, for example $\mathbf{a}, \mathbf{b}, \mathbf{c}$, a basic triangle is defined. Then the affine-invariant coordinates of point $\mathbf{d}$ are determined by the quotient of the area F of the triangles $\mathbf{b}, \mathbf{c}, \mathbf{d}$ or $\mathbf{a}, \mathbf{c}, \mathbf{d}$ and $\mathbf{a}, \mathbf{b}, \mathbf{c}$, respectively. Affine mapping using a transformation matrix T changes the area F of the triangles according to $F' = \det(T)F$. Obviously the quotient of two triangular areas is affine-invariant. Unfortunately this simple procedure has a considerable drawback. For the determination of affine-invariant coordinates of the object and of its image the identical basic triangle has to be used. This equals the need of establishing the correspondence of 3 points of the object and the image. In consequence, the procedure does not fit the aim of real-time identification.

3.1.2 Affine-invariant curvatures

Similar to the standard formulation of the curvature of a planar curve the affine-invariant curvature $\kappa(\tau)$ is defined by

$$\kappa(\tau) = \dot{u}(\tau)\ddot{v}(\tau) - \ddot{u}(\tau)\dot{v}(\tau) \ .$$

With $(u(\tau), v(\tau))$ a parametric formulation of the curve is described as a function of the affine length τ (Naas and Schmid, 1972). As in

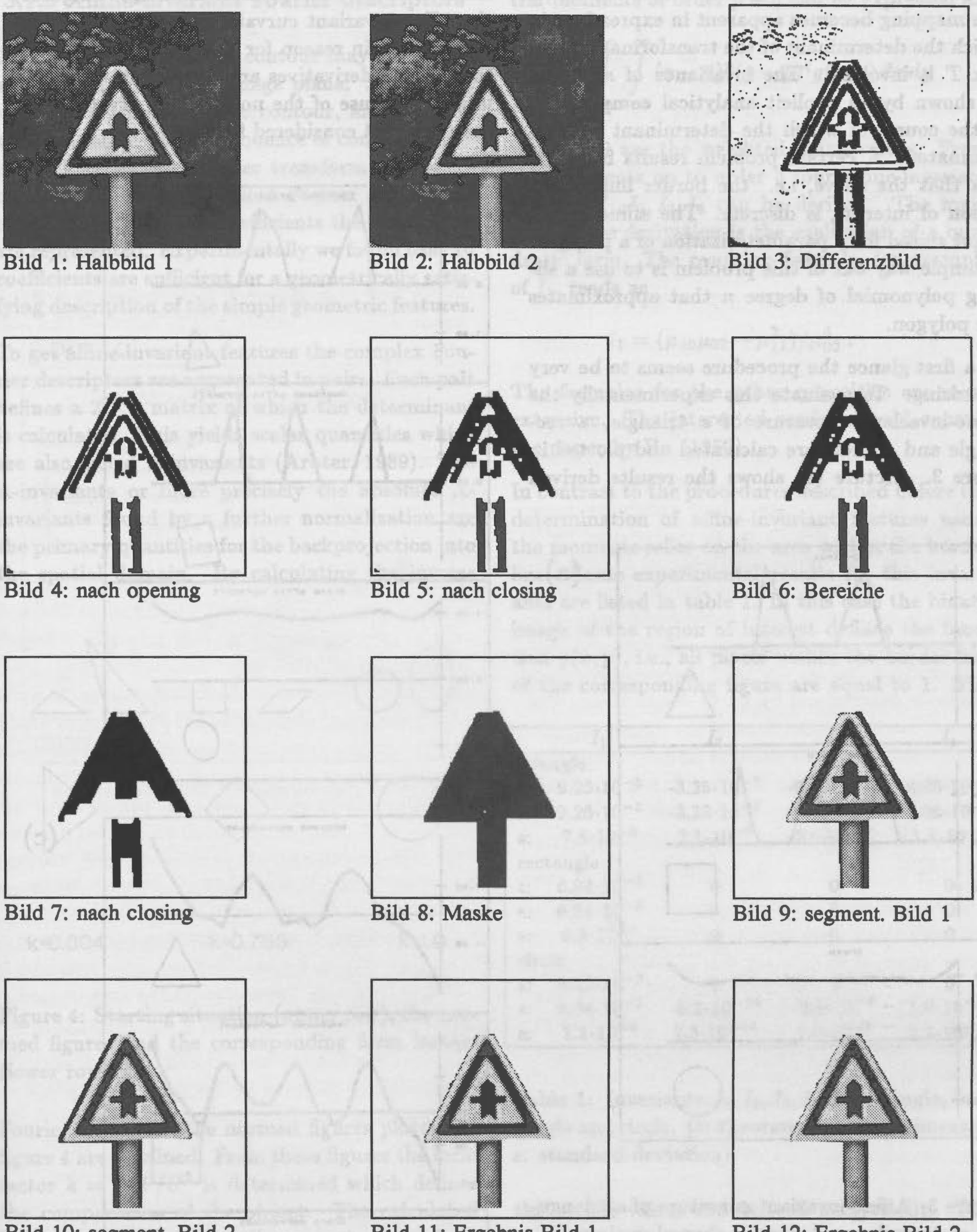


Figure 2: Example: Region of interest extracted by motion based image segmentation

the case of the affine-invariant coordinates an affine mapping becomes apparent in expressions in which the determinant of the transformation matrix T is involved. The invariance of $\kappa(\tau)$ can be shown by an explicit analytical computation in the course of which the determinant $\det T$ is eliminated. A certain problem results from the fact that the curve, i.e. the border line of the region of interest, is discrete. The affine length is not suited for a parameterization of a polygon. A simple way out of this problem is to use a sliding polynomial of degree n that approximates the polygon.

At a first glance the procedure seems to be very promising. To evaluate this experimentally the affine-invariant curvature of a triangle, a rectangle and a circle are calculated and plotted in figure 3. Picture (a) shows the results derived

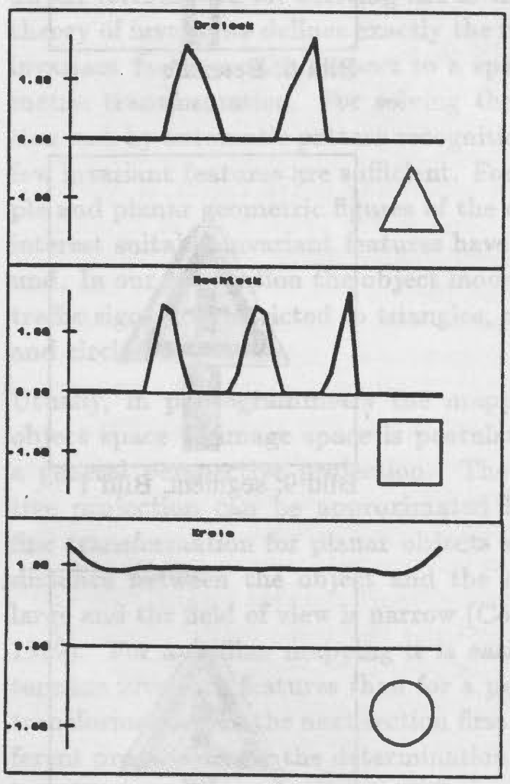
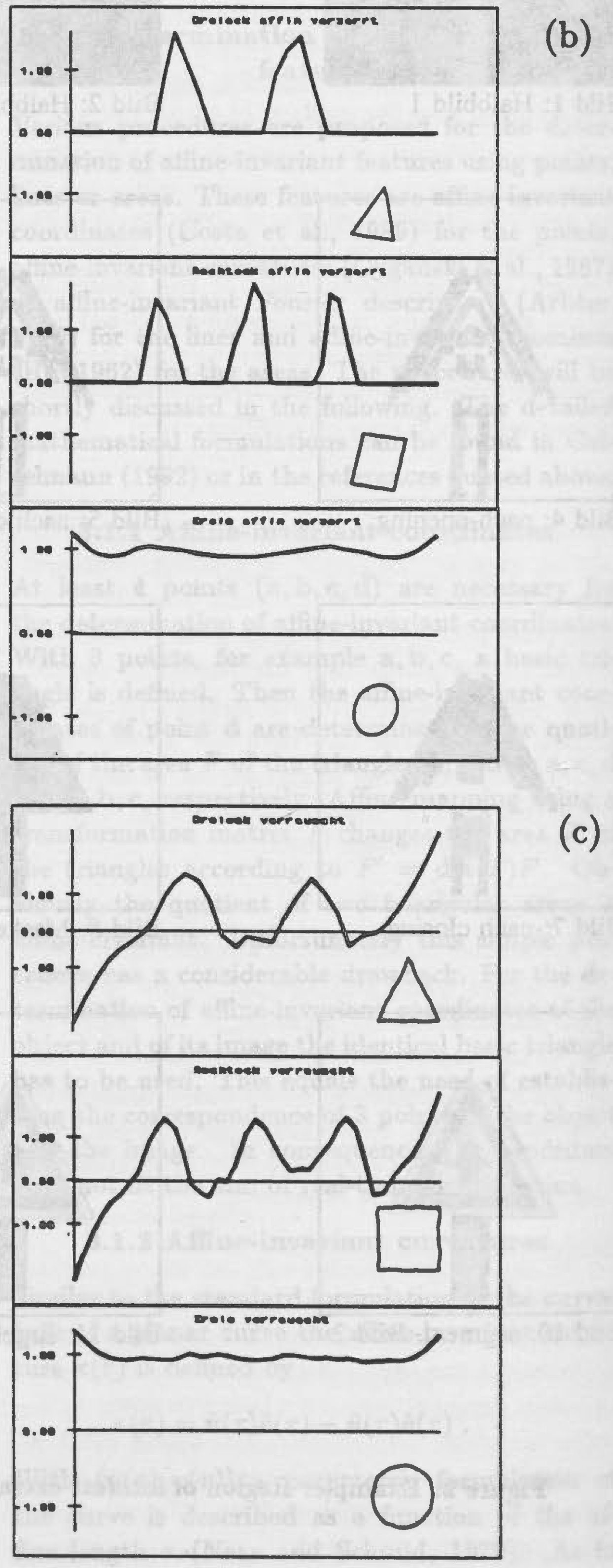


Figure 3: Affine-invariant curvature of a triangular, rectangular and circular shaped contour

from the original figures. In (b) these figures are distorted by affine transformation and in (c) noise is added. The comparison of the plotted curvatures of (c) with respect to (a) and (b) indicates the drawback of the procedure. Though only small noise is added (compare the drawn figures) the affine-invariant curvature changes considerably. The main reason for this noise sensitivity is that second derivatives are involved in the calculation. Because of the noise sensitivity this procedure in not considered further.



3.1.3 Affine-invariant Fourier descriptors

The points of the closed contour may be viewed as being in a complex image plane. Starting at an arbitrary point on the contour, and tracking once around it, yield a sequence of complex numbers. The discrete Fourier transform of this sequence defines the so-called Fourier descriptors of the contour. In this coefficients the frequencies are represented. Experimentally we found that 16 coefficients are sufficient for a geometrically satisfying description of the simple geometric features.

To get affine-invariant features the complex Fourier descriptors are aggregated in pairs. Each pair defines a 2×2 matrix of which the determinant is calculated. This yields scalar quantities which are also called A-invariants (Arbter, 1989). The A-invariants or more precisely the absolute A-invariants found by a further normalization are the primary quantities for the backprojection into the spatial domain. By calculating the inverse

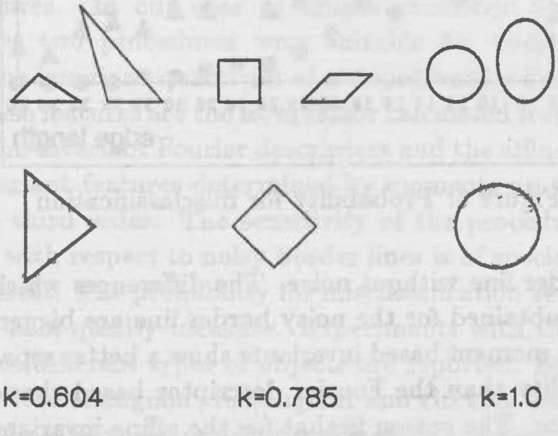


Figure 4: Starting situation (upper row), the normed figures and the corresponding form factors (lower row)

Fourier transform the normed figures plotted in figure 4 are obtained. From these figures the form factor $k = 4\pi F/U^2$ is determined which defines the compactness of the object. The calculated form factors (see figure 4) are a simple feature for the characterization and identification of the unknown object.

3.1.4 Affine-invariant moments

Given a two-dimensional function $g(x, y)$ the central moments of order $p + q$ can be expressed as

$$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q g(x, y)\, dx\, dy$$

where $\bar{x}, \bar{y}$ are the weighted mean values. From this moments up to order 3 four affine-invariant features I_1, I_2, I_3, I_4 can be derived. The main part of the derivation is the evaluation of a quadratic form. The resulting formula, for example of I_1, reads as

$$I_1 = (\mu_{20}\mu_{02} - \mu_{11}^2)/\mu_{00}^4 \,.$$

The formulas for the other quantities are more extensive. The interested reader should consult the paper by Hu (1962).

In contrast to the procedures described before the determination of affine-invariant features using the moments relies on the area within the border line. Some experimental results for this invariants are listed in table 1. In this case the binary image of the region of interest defines the function $g(x, y)$, i.e., all pixels within the border line of the corresponding figure are equal to 1. The

	I_1	I_2	I_3	I_4
triangle				
t:	$9.25{\cdot}10^{-3}$	$-3.25{\cdot}10^{-7}$	$-5.49{\cdot}10^{-5}$	$4.06{\cdot}10^{-6}$
e:	$9.26{\cdot}10^{-3}$	$-3.23{\cdot}10^{-7}$	$-5.47{\cdot}10^{-5}$	$4.06{\cdot}10^{-6}$
s:	$7.8{\cdot}10^{-5}$	$2.1{\cdot}10^{-8}$	$2.0{\cdot}10^{-6}$	$1.8{\cdot}10^{-7}$
rectangle				
t:	$6.94{\cdot}10^{-3}$	0	0	0
e:	$6.94{\cdot}10^{-3}$	0	0	0
s:	$6.8{\cdot}10^{-8}$	0	0	0
circle				
t:	$6.25{\cdot}10^{-3}$	0	0	0
e:	$6.34{\cdot}10^{-3}$	$-8.2{\cdot}10^{-16}$	$-2.1{\cdot}10^{-9}$	$1.0{\cdot}10^{-14}$
s:	$7.1{\cdot}10^{-6}$	$7.3{\cdot}10^{-16}$	$1.0{\cdot}10^{-9}$	$5.2{\cdot}10^{-11}$

Table 1: Invariants I_1, I_2, I_3, I_4 for triangle, rectangle and circle. (t: theoretical, e: experimental, s: standard deviation)

theoretical values are derived analytically. The experimentally found mean values refer to a certain number of calculations with differently affine distorted figures. The empirical values and the theoretical values fit well together. The largest differences can be observed for the circle. In addition to the mean values the corresponding standard deviations are listed. The calculated values refer to triangles, rectangles and circles with an edge size of 20 pixels.

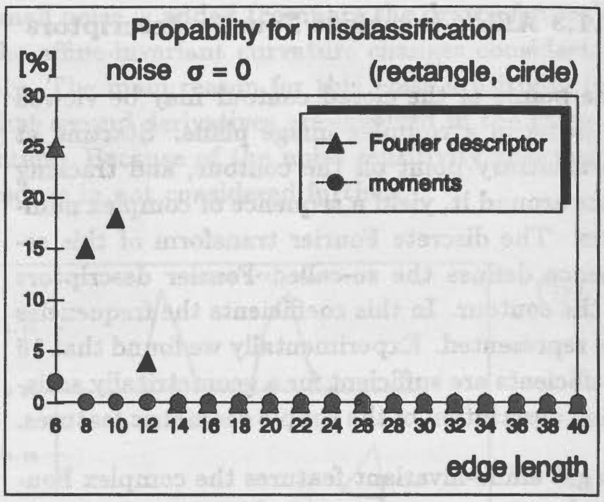
Concerning the separability of rectangle and circle some problems may occur because the only non-zero invariant of both figures is I_1 (cf. table 1). The advantage of this procedure is that the calculation of the moments and invariants is very simple and that no threshold or other parameter has to be defined a priori.

3.2 Empirical investigations for choosing a suitable procedure

The practical suitability of the different procedures described in section 3.1 for classification will now be investigated with the help of simulations. With respect to real-time mapping the procedure should be as simple as possible. On the other hand, the reliability of the procedure must be satisfactory. In this case a high reliability measure is identified with a small probability for a misclassification. With the formulas for the determination of affine-invariant features the probability for a misclassification could be theoretically determined by using the law of error propagation. Because the formulas are very extensive the probability is determined by using simulations. For this purpose the border lines of the figures are differently distorted and noise is added. Then the affine-invariant features are calculated and used to assess the probability for misclassification by the help of the Bhattacharyya distance (Funkunaga, 1972).

For this experiments we restrict on the form factor k which is derived from the affine-invariant Fourier descriptors and on the 4 affine-invariant features I_1, I_2, I_3, I_3 which are derived from the moments up to order 3. Different probabilities for misclassification are received depending on the edge length of the figure and on the amount of noise in the border line. Such a number gives, for example, the probability that a triangle was classified as a circle.

The results show that the largest probability of a misclassification is received for short edges and between the figures rectangle and circle. The diagrams (figure 5) show the probabilities for these both figures. The results of a noise free border line and a noisy border line (Gaussian noise with a standard deviation of 2) are plotted for comparison. Both procedures are equally suitable for a



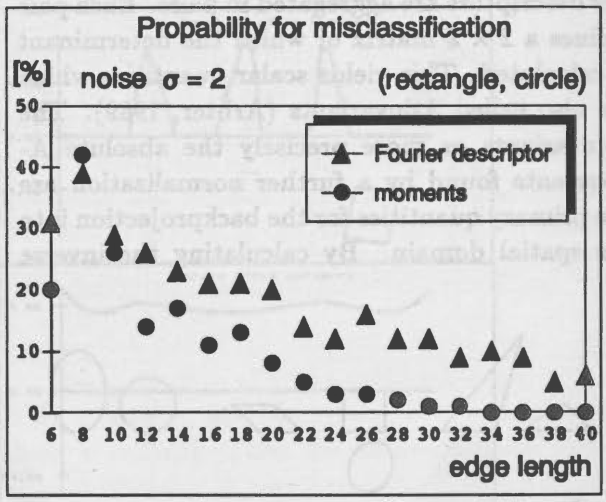


Figure 5: Probability for misclassification

border line without noise. The differences which are obtained for the noisy border line are bigger. The moment based invariants show a better separability than the Fourier descriptor based shape factor. The reason is that for the affine-invariant quantities found by the moments the area is used and not the line information.

As a result one can say that the calculation of the affine-invariant features determined by the moments supplies the best results concerning the separability. Therefore the recognition process should rely on these features. The measured computing time (VAX 3500) for the feature extraction of a 80 × 80 region of interest amounts to 0.2 sec (Fourier descriptor approach) and 0.4 sec (moment approach), i.e. in practice, it is fast enough for real-time mapping applications.

4. CONCLUSION

The aim of the the investigation has been to develop a simple but reliable procedure for the identification and location of simple planar objects. Several general questions were examined that arose during the recognition of simple and planar geometrical figures of an image sequence. Especially the structure of the object and the image motion recorded by an image sequence are exploited to solve the task.

First a procedure was developed for the detection and location of the border line in the image. For objects located in a natural environment the segmentation based on image motion was found to be the most promising procedure. The procedure for extracting the region of interest uses the fact that a change can be identified in a difference image. By that the process corresponds to classical procedures of the analysis of image sequences.

A suitable procedure for reliable object recognition has to be chosen from different existing procedures. In our case of simple geometric features two procedures were suitable for finding affine-invariant quantities of a closed border line. These features are the form factor calculated from affine-invariant Fourier descriptors and the affine-invariant features determined by moments up to the third order. The sensitivity of the procedures with respect to noisy border lines is of special interest. The probability for misclassification serves as a quality measure. Experiments with the three different types of objects are reported. For that the triangular, rectangular and circular border lines of corresponding traffic signs are distorted and noise was added. The determination of the probability for a misclassification applies the maximum likelihood principle. The best results concerning the separability between the three object types are found by determining the affine-invariant features using the moments of an area which lies within a closed border line. Finally the identification of the object is done by maximum likelihood classification. For this purpose the four affine-invariant features can be used directly.

REFERENCES

Arbter, K. [1989]: Affine-invariant Fourier descriptors. In: *From Pixels to Features*, North-Holland, Amsterdam, 1989.

Costa, M., Haralick, R., Phillips, T., Shapiro, L. [1989]: Optimal affine-invariant point matching. In: *Proc. of the DARPA Image Understanding Workshop*, San Mateo, CA, 1989. Morgan Kaufmann.

Cyganski, D., Cott, T.A., Orr, J.A., Dodson, R.J. [1987]: Development, Implementation, Testing, and Application of an Affin Transform Invariant Curvature Function. In: *Proceedings of the 1st International Conference on Computer Vision*, pp. 496-500, London, 1987. IEEE Computer Society Press.

Forsyth, D., Mundy, L., Zisserman, A., Brown, Ch. M. [1990]: Invariance - A New Framework for Vision. In: *Proceedings of the 3rd International Conference on Computer Vision*, pp. 598-605, Osaka, 1990. IEEE Computer Society Press.

Fukunaga, K. [1972]: Introduction to Statistical Pattern Recognition. Academic Press, New York, 1972.

Geiselmann, Ch. [1992]: Zur automatischen Identifizierung einfacher Objekte aus einer Bildfolge. PhD thesis, Universität Stuttgart, Institut für Photogrammetrie, Stuttgart, 1992.

Hu, M.-K. [1962]: Visual Pattern Recognition by Moment Invariants. *IRE Transactions on Information Theory*, pp. 179-187, 1981.

Jain, R. [1981]: Extraction of Motion Information from Peripheral Processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 3, no. 5, 1981.

Naas, J., Schmid, H.L. [1972]: Mathematisches Wörterbuch. Akademie-Verlag, Berlin, Teubner Verlagsgesellschaft, Stuttgart, 1972.

Novak, K. [1990]: Integration of a GPS-receiver and a stereo-vision system in a vehicle. In: *International Archives of Photogrammetry and Remote Sensing*, vol. 28, 5/1, pp. 16-23, Zürich, 1990.

Schwarz, K.P. [1992]: Inertial Techniques in Geodesy - State of the Art and Trends. In: *High Precision Navigation 91*, Linkwitz, K., Hangleiter, U. (ed.), Dümmler Verlag, Bonn, 1992.

Identification et repérage d'objets simples en cartographie en temps réel

Résumé

Dans le présent exposé, on se penchera sur deux aspects de la cartographie en temps réel. D'abord, on traitera de la combinaison des capteurs de collecte de données, comme les capteurs de positionnement et d'imagerie utilisés en cartographie par ordinateur. Ensuite, on considérera l'élaboration d'une méthodologie cartographique, c'est-à-dire de méthodes et d'algorithmes d'identification et de repérage d'objets sur lesquels des renseignements doivent être captés et stockés dans une base de données. Étant donné l'aspect temps réel de ce type de cartographie, ces méthodes doivent être simples et fiables. En pratique, cependant, il en est tout autrement. Une pièce d'équipement cartographique permet à son *utilisateur* d'identifier et de repérer des objets dans un cadre de référence absolu.

Notre exposé portera sur une opération de repérage et d'identification des panneaux de signalisation routière. Nous avons élaboré pour cette opération une méthode permettant de réduire considérablement le travail du cartographe dans la préparation d'une carte des panneaux de signalisation. Cette méthode de cartographie automatique comporte plusieurs étapes. On segmente d'abord l'image en une zone d'intérêt et une zone d'arrière-plan à l'aide d'un couple stéréoscopique révélant le mouvement. On dérive ensuite le champ vectoriel de décalage pour cette région. La déconvolution du flou attribuable au mouvement produit une meilleure segmentation. On détermine ensuite le contour clos de la zone d'intérêt et on dérive ses caractéristiques affines et invariantes. Enfin, on identifie des classes d'objets pour les panneaux de signalisation repérés en déterminant la classification la plus probable à partir de ces caractéristiques. Les expériences décrites dans cet exposé ont été réalisées en laboratoire et sur le terrain.

LASER RANGE SCANNER SUPPORTING 3-D RANGE AND 2-D GREY LEVEL IMAGES FOR TUNNEL SURFACE INSPECTION

C. Fröhlich, M. Mettenleiter and G. Schmidt
Laboratory of Automatic Control Engineering
Technical University of Munich, D-80290 Munich, Germany

KEY WORDS: range sensor, range finders, cw lasers, range data, integration, tunnel inspection

ABSTRACT

For the survey and inspection of buildings, tunnels or underground canals, a non-tactile, robust and precise imaging of height and depth profiles is a basis sensor technology. For visual inspection, surface classification, and documentation purposes, however, additional information concerning reflectance of measured objects is necessary. High-speed acquisition of both geometric and visual information is achieved by means of an eyesafe laser range scanner, supporting consistent range and grey level profiles which are similar to video images.

This paper reports the optical, electronic and mechanical system design of the scanner, its information processing system including dedicated signal processors, and a transputer network for sensor data processing. The paper discusses experimental results of the laser range scanner with respect to noise as well as its long-term stability, and accuracy. Finally, results from tunnel surface inspection with the laser range scanner are presented.

1. INTRODUCTION

Range sensing is a crucial component of automation in industrial processes. To realize robotic systems [1,2] for manufacturing, handling, transportation and inspection tasks, however, it is the only way to provide the system with three dimensional information [3] of its environment. The classical computer vision approach to range sensing is to use passive techniques such as stereo-vision or motion stereo. However, those techniques are not yet sufficiently reliable or fast enough to be used in many applications, most notably real-time robotic or inspection systems. Active sensors, which generate the illumination instead of using only the ambient illumination, have received increasing attention as a viable alternative to passive sensors, featuring direct access to range information in real-time. Their importance was recognized relatively early. For example, Besl [4] examines a wide variety of active range imaging technologies, comparing them quantitatively by evaluating a figure of merit based on range accuracy, field of view, and image acquisition time. A review of using range sensing devices in autonomous navigation of mobile robots can be found in Hebert ([5], Everett [6], Elfes [7] and Rozmann[8]. Nitzan et al. [9] describe a system interpreting indoor scenes by use of range and intensity information from a laser ranging system.

Through the framework of SFB 331 "Information Processing in Autonomous Mobile Robots", a 3-D laser range camera [10,11,12] was developed for indoor vehicle and robotic applications. The laser range camera is an optical-wavelength radar system, and is comparable to devices built by Erim [13], Odetics [14], and Perceptron [15], measuring the range between sensor and target surface as well as the reflectance of the target surface which corresponds to the magnitude of the back scattered laser energy. In contrast to these range sensing devices, the laser camera under consideration is designed for eyesafe operation, emitting a minimum of near-infrared laser energy.

For surface inspection of railway and highway tunnels however, the performance of the range measuring system of the laser camera was improved with a look towards robustness, range accuracy, and range precision.

This paper reports design and practical details of the laser range scanner for tunnel surface inspection. Chapter 2 outlines the performance requirements with tunnel surface inspection and introduces the measurement principle of the sensor. Hardware design of the laser range scanner, including the main modules, such as the laser head with transmitter and receiver electronics, high frequency unit, laser beam deflection system, and digital signal processing unit are discussed in Chapter 3. The signal processing unit consists of dedicated signal processors and a transputer system for real-time sensor data processing. Chapter 4 focuses on statistical evaluation of range data, including noise, drift over time, precision, and accuracy with range measurements. It discusses the influences of ambient light, surface material of the target, and ambient temperature for range accuracy and range precision. Furthermore, experimental results from the inspection of tunnel surfaces are presented. Chapter 5 concludes the paper by summarizing its results and gives a short outlook to future work in the tunnel surface inspection project.

2. DESIGN ASPECTS

2.1 Performance requirements

With railway tunnel inspection, the precise measurement of the tunnel tube geometry, tracks and contact wires (Fig. 3a) is essential in order to hold tolerances for safe passing of trains. Furthermore, detection of possible cracks (even capillary cracks) in the tunnel surface is important in order to initiate repair work of the tunnel tube in time. For tunnel inspection, measurement of 360° profiles while navigating through the tunnel tube is more sufficient than measuring 3-D images. To meet these demands, high spatial resolution, accurate and highly reliable range measurements are necessary for geometric inspection of tunnel tubes.

Gap-less inspection of tunnel tubes requires a range of up to 10 m, a spatial resolution of 2500 pixels per 360° profile, and a distance between two consecutive measured profiles of less than 2.5 cm. Due to sooty walls and metallic objects in the tunnel, the sensor system has to handle high dynamic range reflectances of the objects. Furthermore, the sensor system must be robust when dealing with environmental influences such as temperature or humidity as well as varying illumination conditions (dark, ambient light, lamps, etc) as they are typical for tunnel tubes. Safe operation with respect to people is required all times.

Only non-tactile sensors are suited for covering these demands. Non-tactile range measurement techniques may be classified as either active techniques, directing visible or infrared (IR) light, ultrasonic [16] or radar [17] pulses to the surface to be measured, or as passive techniques based on vision. A rich variety of passive vision techniques produce three-dimensional information. Traditionally they have lacked robustness with changing illumination conditions and generality, and have not proven themselves effective in practice. Passive stereo or motion stereo vision [18,19] are particullary promising sources of range information, but require substantial data processing to match images with each other in order to determine range by triangulation and, therefore, are not well suited for real-time tunnel surface inspection.

For these reasons we have selected an active range measurement technique which directly determines *"range"* data with a minimum computation time. To achieve high spatial resolution, only an active technique emitting collimated laser light is suitable. The collimated laser beam is directed to the target to be measured and the back scattered light is sensed. In addition to *"range"* measurement, evaluation of the magnitude of the back-scattered light provides an *"active grey level"* which is similar to the grey level information of a video camera. Both range and grey level data of a target point are registered at the same time and correspond to a single target point defined by the laser beam direction. Due to emitted laser energy, both informations, range, and grey level data are nearly independent from environmental influences and illumination conditions.

In order to achieve profile data the range measuring system is combined with a one-dimensional scanner for 360° beam deflection. For longitudinal inspection of the tube, the system is mounted on a special vehicle (or train) navigating at a maximum speed of 5 m/s through the tunnel. Resulting spiral profiles are combined with respect to the corresponding sensor positions. The final range image reflects geometric dimensions of the tunnel tube whereas the grey level image is used for visual inspection, surface classification, and documentation purposes.

2.2 The two-frequency phase-shift method

Because of the requirements for high-precision range measurement within a range of up to 10 meters, evaluation of the phase-shift (Fig. 1) between a reference laser beam and the back scattered laser light is more suitable than measuring the light's extremely short time of flight.

The amplitude P_E of the emitted continuous-wave laser signal is simultaneously intensity modulated (am-cw) with two frequencies ω_1 and ω_2. Laser light back scattered from a target is collected by an avalanche photodiode. The amplitudes P_D are fairly small. Due to the time-of-flight, the received sine-shaped signals are phase-shifted in relation to their reference in the transmitted signal. Phase shifts φ_1 and φ_2 are proportional to the range d and the modulation frequencies. Since phase shifts are only unique modulo 2π, the modulation frequencies ω_1 and ω_2 are selected to provide a sufficient measurement range with an appropriate range resolution. A low-frequency signal (LFS: ω_1 = 10 MHz) guarantees a coarse but absolute measurement range of s_1 (s_1 = 15 m), whereas the high-frequency component (HFS: ω_2 = 80 MHz) provides a fine (s_2 = 1.875 m) but ambiguous range information over s_1. Correct combination of the phase shifts φ_1 and φ_2 of both frequencies provides absolute and accurate range measurements within the specified range.

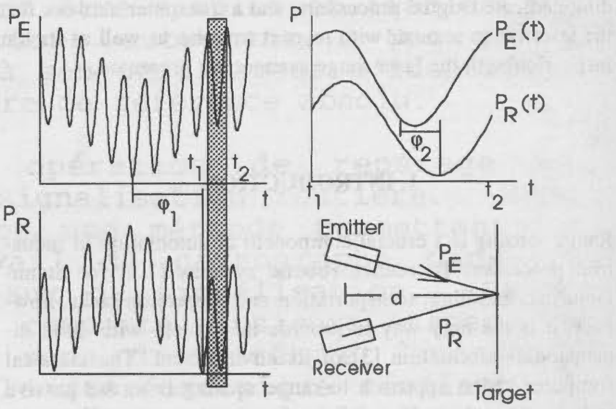


Fig. 1: The two-frequency phase-shift method

3. THE LASER RANGE SCANNER

Hardware of the laser range scanner (Fig. 2) consists of two major components: the range measuring system and the beam deflection system. Both components operate independently from each other. They are connected via a control and monitoring system.

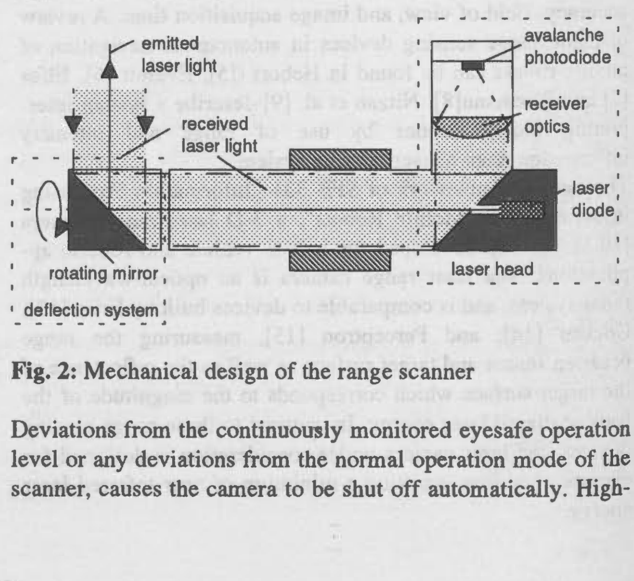


Fig. 2: Mechanical design of the range scanner

Deviations from the continuously monitored eyesafe operation level or any deviations from the normal operation mode of the scanner, causes the camera to be shut off automatically. High-speed data transfer between the laser scanner and a sensor data computer is achieved by a transputer link (20 Mbit/s).

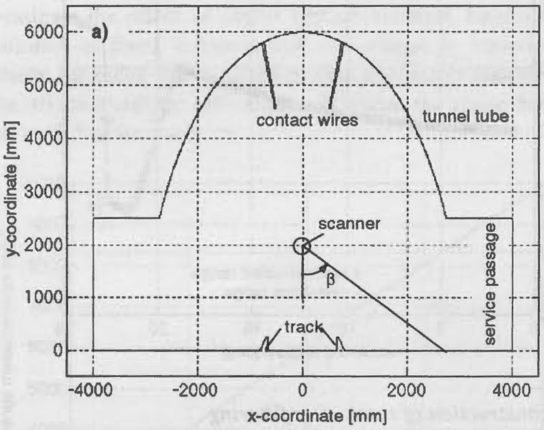
3.1 Range measuring system

The range measuring system is the major component of the laser range scanner. It consists of the laser head, high frequency unit and the signal processing unit for range data preprocessing.

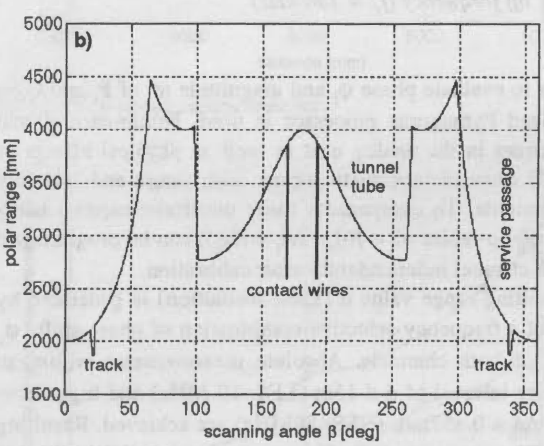
Laser Head:
The laser head emits a continuous wave, intensity modulated infrared (IR) laser beam and detects the laser light back-scattered from a target. It consists of several electro-optical and micro-mechanical components.
The laser beam is generated by a semiconductor laser diode that emits 20 mW of laser power at 810nm (near infrared). Varying the drive current to the diode (±15 mW_{PP}, am-cw) modulates the amplitude of the laser light. Eyesafety (DIN EN 60825) is achieved by reducing emitted laser power by means of a grey-glass filter to P_E = 4.5mW (laser class 3A). Optical power reduction minimizes nonlinear effects of the laser diode as well as temperature and noise effects. A set of microlenses forms the laser cone from the diode into a coaxial path of rays with small divergence (δ = 0.01 mrad). The laser light (P_R) reflected from an environmental object returns on the same optical path to the receiver where receiver optics (Ø = 60 mm, f = 50 mm) focus it onto the detector. Using an avalanche photodiode as a detector and an IR filter for elimination of spectral noise in laser light guarantees high dynamic range with the reflectance (P_R/P_E = 2% .. 99%) of environmental objects.

High frequency unit:
The high frequency unit generates a modulation signal for the laser diode and consists of two quadrature receivers for frequency-selective demodulation of the back-scattered laser light.
For modulation of the emitted laser light two sine signals with the frequencies ω_1 = 10 MHz and ω_2 = 80 MHz are necessary. In order to achieve in-phase correlation of the two signals, a single 80 MHz oscillator in combination with a divider generates modulation frequencies. The resulting signal for drive current modulation (Fig. 1) of the laser diode is generated by eliminating harmonic signals in the square-wave signals of the oscillator and adding sine signals.
The back-scattered laser light is amplified, and band-pass filtered in order to seperate LFS (10 MHz) and HFS (80 MHz) channels, and then downconverted. Using homodyne quadrature mixers for distortion-free demodulation guarantees high dynamic range with detected signal levels (S/N ≈ 90 dB). The quadrature mixers evaluate the respective in-phase (P_i) and quadrature (Q_i) signal of each channel.
In-phase (P_i) and quadrature signal (Q_i) of each channel form a complex measurement vector. Phase φ_i of the vector is proportional to measured *"range d_i"* in the channel's specific ambiguity interval, whereas magnitude m_i of the vector directly indicates intensity *"active grey level g_i"* of back scattered light.

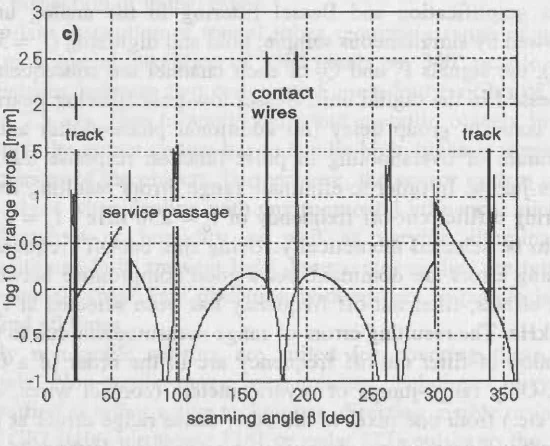
Signal processing unit:
After amplification and Bessel filtering in the analog unit, followed by simultaneous sample, hold and digitizing (f_s = 500 kHz), the signals P_i and Q_i of each channel are subsequently processed in the digital unit. Bessel low-pass filtering guarantees constant group delay (no additional phase-shifts) and a minimum of overshooting in pulse-function response due to range-jumps. In order to eliminate range errors resulting from filtering a filter cut-off frequency of f_c = 250 kHz ($f_c = f_s/2$) has to be selected theoretically. Using this cut-off frequency, aliasing errors are dominant. As a good compromise between both effects, filter cut-off frequency has been selected at f_c = 130 kHz. The resulting errors of range measurement due to the adaption of filter cut-off frequency are in the order of a few mm. Only range-jumps of several meters (contact wires, edges, etc.) from one pixel to the other cause range errors at the respective jump-edges of several cm (Fig. 3). Hybrid 14-bit A/D converters are used to digitize filtered signals P_i and Q_i of each channel in order to fulfill the demands of high dynamic range (S/N ≈ 56 dB) with reflectance of target surfaces and the high accuracy (S/N_{min} ≥ 25 dB) of range measurements.



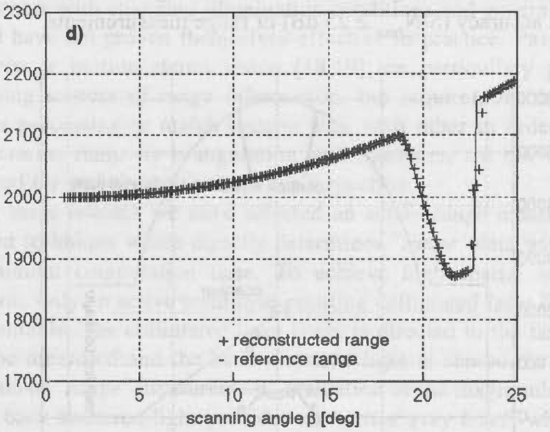
3a) Cartesian plot of reference tunnel



3b) Polar plot of reference tunnel profile



3c) Range errors resulting from filtering



3d) Reconstruction of track after filtering

Fig. 3: *Range errors depending on filter cut-off frequency (f_c = 130 kHz)*

In order to evaluate phase φ_i and magnitude m_i of P_i and Q_i, a specialized Pythagoras processor is used. Nonlinearities and offset errors in the analog unit as well as physical effects of the APD cause deterministic errors with range and intensity measurements. To compensate these uncertainties, two hardware look-up-tables ($d_i = f(d_i)$; $\Delta d_i = f(g_i)$) can be programmed for each channel independently after calibration.

The resulting range value d (15bit resolution) is generated by means of a frequency-selective combination of phase shifts φ_1 and φ_2 of both channels. Absolute measurements within an ambiguity interval of $d \leq 15$m (LFS: 10 MHz) and high accuracy of $\Delta d = 0.457$mm (HFS: 80MHz) are achieved. Resulting grey level g corresponds to the magnitude of the HFS with 12-bit resolution.

Pixel oriented range data preprocessing, i.e. filtering software, coordinate transformation, and classification of single pixels as well as image processing algorithms are implemented on a transputer network based on seven T805 transputers.

3.2 Deflection system

For surface inspection of railway and highway tunnels, 360° profile measurement is required. Scanning of 360° profiles is achieved through the deflection of the emitted laser beam over a high-speed rotating mirror (Fig. 2). In order to fulfill the demand for gap-less tunnel inspection in combination with a speed of up to 5 m/s of the carrier vehicle (i.e distance between two consequtive profiles is less than 2.5 cm) the rotation speed of the mirror is controlled to 200 rps.

Table 1 gives a survey of currently achieved system parameters.

Resulting spiral profiles of the tunnel profile scanner are combined with respect to the actual sensor position. The final *range image* reflects geometric dimensions of the tunnel tube whereas the *grey level image* is used for visual inspection, surface classification, and documentation purposes.

Laser head	
IR semiconductor laser diode (λ = 810 nm)	
Hybrid APD photosensor	
Emitted laser power:	4.5 mW
High frequency unit	
Two-frequency phase-shift method	
Modulation frequencies:	10/80 MHz
Signal processing unit	
Reflectance (P_R/P_E):	2% ... 99%
Depth of field:	0 ...15 m
Accuracy of range data:	up to 0.45 mm
Sampling rate:	500 kHz
Scanner system	
Beam deflection through rotating mirror	
Field of view:	360° profiles
Spatial resolution:	2500 pixel/profile
Profile imaging rate:	200 profile/s

Tab.1: *System parameters*

4. EXPERIMENTAL RESULTS

In this section we describe a series of experiments designed to measure accuracy of the laser range scanner under different conditions and to compare it with the predicted theoretical values. Following Besl [4], *accuracy*, the difference between measured range and true range, and *precision*, the variation of measured range to a given target, are distinguished.

We conduct experiments within structured test scenes by means of special test objects with standardized reflectances R (R ∈ [2%, 5%, 20%, 50%, 75%, 99%]) in our laboratory. The area has both natural illumination that enters through windows, and artificial illumination generated by lamps mounted on the ceiling. These lamps will be referred to as *spotlights* because they are powerful and somewhat directional.

4.1 Accuracy

To determine the accuracy of the range measurements is to identify the deviation between measured range and true range.

For a target point, let *d* be the range measurement reported by the scanner, and let *r* be the true range from the geometric origin of the scanner, measured with a high precision reference system. Under ideal conditions, we expect to observe that *d* = *r*, but nonlinear effects with range measurements, offset errors in the analog unit, surface material of the target, and changing lighting conditions cause errors with range measurement.

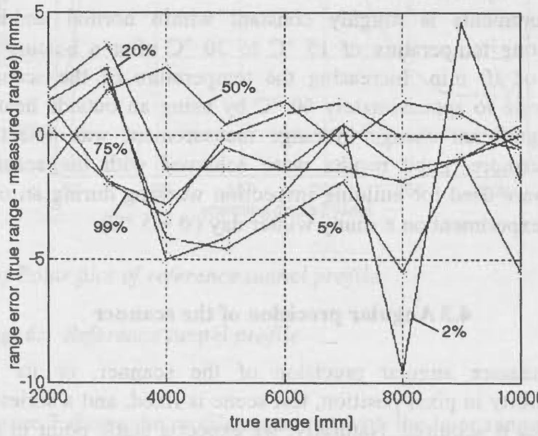


Fig. 4: Sensor accuracy for different materials

To determine range accuracy we acquired range measurements of targets with various surface materials and reflectance within the ambiguity interval of the scanner. The target is positioned at known distances with an angle of incidence α of about 90°, i.e. normal to the target surface. For each target position (0 ... 15 m; 0.5 m steps) with all standardized reflectances R, a sample of j = 1000 consecutive range measurements is acquired. Experimental results for linearity, and worst-case deviations e = max (e_j | $e_j = d_j - r$) between measured range and reference range r are illustrated in Fig. 4. After calibrating the look-up-tables $d_{10/80} = f(d_{10/80})$ of both measurement channels, high accuracy with linearity (max. error 0.1 % of max. range) and resulting range errors of less than ±10 mm for worst-case reflectance of R = 2 % are achieved.

Range drift

To measure range drift, a single target is placed (d = 6m; α = 90°; R = 50%) in front of the laser scanner (ϑ = 18°C) and range data is acquired over several hours. Due to thermal heating effects in the measuring system (laser head and high frequency unit) a drift of 4 cm in range measurement occured in the first 10 min. After this "heating time", no significant range drift (Δd ≤ e) can be detected for hours. Compared [20] to other laser range sensors, the laser range measuring system shows good long-time stability and high accuracy with range measurements.

4.2 Precision

For characterizing the precision of the range measurements the variation of repeated range measurements is identified. Ambient lighting conditions, surface material of the target, distance from the scanner to the target, beam incidence angle at the target, and ambient temperature are varied in this paper. Under each of these conditions 1000 range measurements are sampled at each target and each position. Mean and variance of the depth measurements are computed, and final precision is quantified as the standard deviation of the distribution of the measurements.

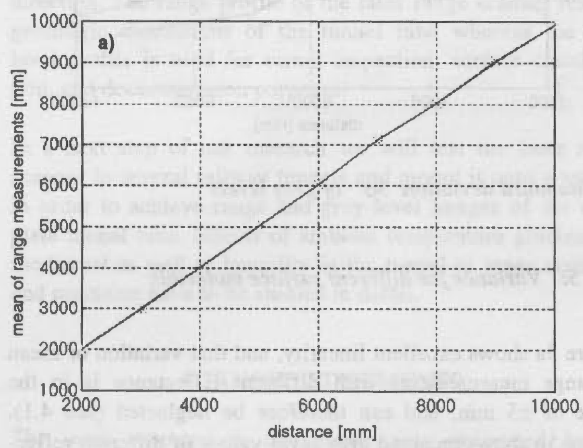
Effect of ambient light

To study the effect of ambient light on sensed range, a black target (R = 5%) is placed at a fixed distance, samples are taken, and mean and variance are computed. This procedure is repeated for target positions between 2 m and 10 m (2 m interval) in front of the sensor under different indoor lighting conditions: during night, during a sunny day, and during a cloudy day, all with and without spotlights.

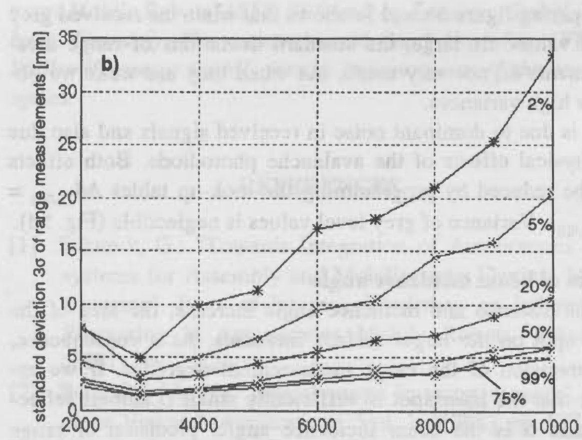
Evaluation of experiments shows that precision is nearly independent from intensity of ambient illumination. Only when sunlight or a spotlight directly radiates into the scanner, noise with range measurement increases, therefore decreasing precision with measurements.

Effect of surface material

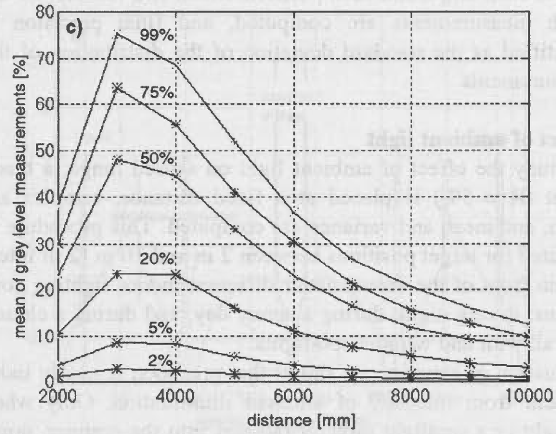
Surface material and therefore the reflectance of a target surface has a definite effect on the range measurements. To investigate the effect of object surface material, targets are positioned at fixed distances and reflectance is varied. Fig. 5 shows the mean and standard deviation of range measurements for all standardized reflectances R within the range from 2 to 10 m (0.5 m interval).



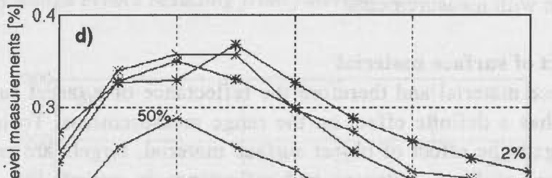
5a) Mean of range measurements



5b) Standard deviation 3σ of range



5c) Mean of grey levels



5d) Standard deviation 3σ of grey levels

Fig. 5: Variance for different surface materials

Figure 5a shows excellent linearity, and that variation of mean of range measurements with different reflectance is in the range of ±5 mm, and can therefore be neglected (see 4.1). Figure 5b shows received grey level values of different reflectances over the whole ambiguity interval of the laser range scanner and confirms theoretical considerations.

Comparing figure 5b and 5c shows that when the received grey level values are large, the standard deviations of range measurements do not vary much, but when they are weak we observe high variances.

This is due to dominant noise in received signals and also due to physical effects of the avalanche photodiode. Both effects can be reduced by programming the look-up tables $\Delta d_{10/80} = f(g_{10/80})$. Variance of grey level values is neglectible (Fig. 5d).

Effect of beam incidence angle

As the distance and incidence angle increase, the area of the laser spot on the target surface increases. As a consequence, the precision of the range measurement degrades. If we assume that the laser spot is sufficiently small (Lambert reflector) and α is the beam incidence angle, precision of range measurement is as a first order approximation inversely proportional to $\cos(\alpha)$.

Experiments by rotating a target (R = 50%) in order to produce different incidence angles confirm that dependence, indicating it as reasonable approximation for the the laser range measuring system for $20° \leq \alpha \leq 90°$ within the whole field of view.

Effect of ambient temperature

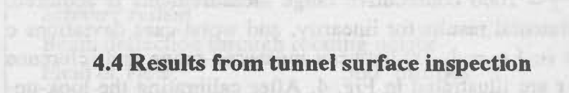
Experiments demonstrated that precision of the laser range scanner by means of mean and standard deviation of range measurements is roughly constant within normal ambient operating temperature of 15 °C to 30 °C after a heating up time of 10 min. Increasing the temperature of the scanner hardware to approximately 60 °C by using an outside heater, no significant change in range measurement was detected. Furthermore, good results were achieved with the scanner hardware used for building inspection working during an outdoor experiment on a sunny winter day ($\vartheta \approx 5$ °C).

4.3 Angular precision of the scanner

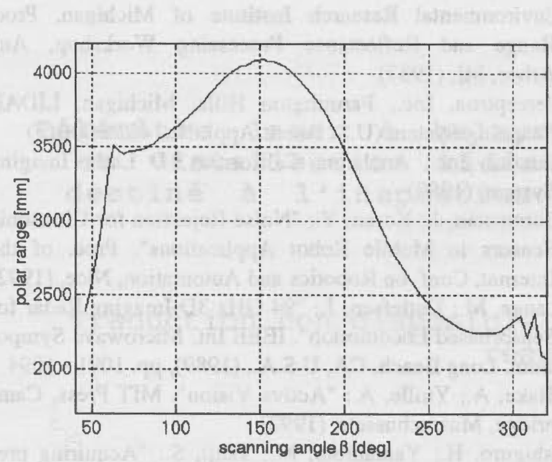
To measure angular precision of the scanner, or its repeatability in pixel position, test scene is fixed, and a series of profiles is acquired. Naturally, we expect a static point in the scene to project to one and only one pixel position in each of the profiles. The scanner is positioned in front of a small wall (d = 5 m; width = 500 mm). Evaluation of both range data and object width by means of standard deviation of jump edges of the wall suggest that the angular precision of the scanner is not a limiting factor for the system. But due to the scanning mechanism with a resolution of 2500 points per profile, only objects with a minimum width of 3 pixels, i.e. about 8 mm/m, can be reliably detected.

4.4 Results from tunnel surface inspection

To examine tunnel surface inspection by means of the laser range scanner, geometric dimensions are evaluated. Figure 6 shows the reference profile of a real tunnel tube measured in Switzerland. All geometric dimensions of the tunnel tube are known by means of a reference range sensor system. Tunnel surface reflectance (sooty and black tube surfaces) is in the order of R = 2% to R = 20%.



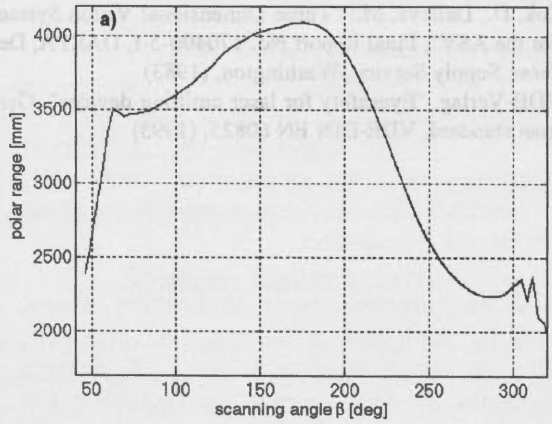
6a) Cartesian plot of reference tunnel profile



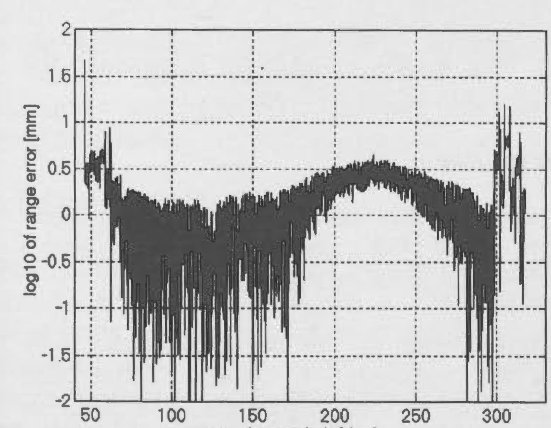
6b) Polar plot of reference tunnel profile

Fig. 6: *Reference tunnel profile*

Figure 7 shows the profile measured with the laser range scanner and compares it to the reference tunnel.



7a) Measured profile (polar plot)



7b) Deviations with profile measurement

Fig. 7: *Tunnel profile measurement results*

Resulting deviations between reference and measured profile of several mm correspond to examinations of Chapter 3 and 4 and fulfill all demands with tunnel surface inspection. Corresponding grey level profile is used for visual inspection, surface classification, and documentation purposes.

5. CONCLUSIONS

In this paper we have examined in detail the measurements supplied by an amplitude-modulated laser range sensor. We presented experimental performance data and discussed the application of the sensor hardware for tunnel inspection. The quantitative performance, in terms of accuracy and precision with different surface material types, fulfills the demands of tunnel surface inspection. Compared [20] to other laser range sensors, the laser range measuring system shows good performance data. Robustness of the laser range scanner even under varying temperature and lighting conditions is sufficient for application in industrial environments. Ambiguity interval of range measurement and range accuracy of the laser range scanner may be adapted to other applications by selecting other modulation frequencies.

Parallel to *"range"* measurements reflectance of surface material *"active grey level"*, is measured. Both, range and grey level data of a target point are registered at the same time and correspond to one single target point defined by the laser beam direction. The range profile of the laser range scanner reflects geometric dimensions of the tunnel tube whereas the grey level profile is used for visual inspection, surface classification, and documentation purposes.

In a next step of our research we will test the laser range scanner in several railway tunnels and mount it onto a vehicle in order to achieve range and grey-level images of the complete tunnel tube. Effects of ambient temperature gradients in the tunnel as well as humidity in the tunnel to range accuracy and precision have to be studied in detail.

ACKNOWLEDGEMENTS

The work reported in this paper was supported by the Deutsche Forschungsgemeinschaft, as part of an interdisciplinary research project on "Information processing in Autonomous Mobile Robots" (SFB 331) and by Spacetec GmbH, Freiburg, Germany. The authors would thank Dipl.-Ing. (FH) R. Wißler (Spacetec GmbH) for the improvements of the receiver optics.

REFERENCES

[1] Schmidt, G.: "Towards Integration of Autonomous Subsystems for Assembly and Mobility into Flexible Manufacturing". Proc. of Internat. Workshop on Information Processing in Autonomous Mobile Robots, München, (1991), pp. 3 - 20

[2] Besl, P.: "Active, optical range imaging sensors". Machine Vision & Applications. Vol. 1, (1988), pp. 127-152

[3] Kanade, T.: Three-Dimensional Machine Vision. Kluwer Academic Publishers, 1987

[4] Besl, P.: "3-D range imaging sensors". Technical Report GMR-6090, General Motors Research Lab., Warren, MI, USA, (1988)

[5] Hebert, Krotkov, E.; M.; Kanade, T.: "A Perception system for a planetary explorer". Proc. of IEEE Conference on Decision and Control, Tampa, FL, (1989), pp. 1151 - 1156

[6] Everett, H.: "Survey of collision avoidance and ranging sensors for mobile robots". Robotics & Autonomous Systems, Vol. 5, (1989), pp. 5 - 67

[7] Elfes, A.: "A stochastic spatial representation for Active Robot Perception". Proc. on the sixth Conference on Uncertainty and AI, AAAI, Cambridge, MA, (1990)

[8] Rozmann, M.; Detlefsen, J.: "Environmental Exploration based on a 3-D Imaging Radar Sensor". Proc. of the Internat. Conf. on Intelligent Robots and Systems, Raleigh, (1992), pp. 422 - 429

[9] Nitzan, D.: "Assessment of robotic sensors". Proc. of International Conference Robot Vision and Sensory Controls, London, UK, (1981), pp. 1 - 11

[10] Karl, G.: "Eine 3-D Laser-Entfernungskamera zur Bewegungsführung mobiler Roboter". Laboratory of Automatic Control Engineering, Dissertation, Technical University of Munich, (1990)

[11] Fröhlich, C.; Schmidt, G.: "A Three Dimensional Laser Range Camera for Sensing the Environment of a Mobile Robot". Sensors and Actuators A, 25 - 27, (1991), pp. 453 - 458

[12] Fröhlich, C.; Schmidt, G.: "Laser-Entfernungskamera zur schnellen Abbildung von 3-D Höhen- und Tiefenprofilen". Sensoren - Technologie und Anwendung, GMA, Bad Nauheim, (1994), pp. 543 - 549

[13] Environmental Research Institute of Michigan, Proc. Range and Reflectance Processing Workshop, Ann Arbor, MI, (1987)

[14] Perceptron, Inc., Farmington Hills, Michigan, LIDAR Scanning System (U.S. Patent Appl. No. 4226-00015)

[15] Odetics, Inc., Anaheim, California 3-D Laser Imaging System, (1989)

[16] Borenstein, J.; Koren, Y.: "Noise Rejection for Ultrasonic Sensors in Mobile Robot Applications". Proc. of the Internat. Conf. on Robotics and Automation, Nice, (1992)

[17] Lange, M.; Detlefsen, J.: "94 GHz 3D-Imaging Radar for Sensorbased Locomotion". IEEE Int. Microwave Symposium, Long Beach, CA, U.S.A., (1989), pp. 1091 - 1094

[18] Blake, A.; Yuille, A.: "Active Vision". MIT Press, Cambridge, Massachussets, (1992)

[19] Ishiguro, H.; Yamamoto, M.; Tsuji, S.: "Acquiring precise range information from camera motion". Proc. of Internat. Conf. on Robotics and Automation, Sacramento, (1991), pp. 2300 - 2305

[20] Hebert, M.; Krotkov, E.: 3D measurements from imaging laser radars: how good are they?". Image and Vision Computing Magazine, vol 10 no 3 (1992), pp. 170 - 178

[21] Kweon, I.: "Modelling Rugged Terrain with Multiple Sensors",.PhD thesis, Scool of Computer Science, Carnegie Mellon University (1991)

[22] Zuk, D.; Delleva, M.: "Three-Dimensional Vision System for the ASV". Final Report No. 170400-3-f, DARPA, Defense Supply Service, Washington, (1983)

[23] VDE-Verlag, "Eyesafety for laser emitting devices". German standard, VDE-DIN EN 60825, (1993)

Télémètre laser à balayage pour mesure tridimensionnelle et visualisation bidimensionnelle en niveaux de gris destiné à l'inspection des canalisations souterraines

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif

The Accuracy of Features Positioned with the GPSVan

Guangping He[1], Kurt Novak[2], Wei Tang[2]

[1]Center for Mapping, The Ohio State University
1216 Kinnear Rd, Columbus, OH-43212, USA

[2]Department of Geodetic Science and Surveying,
The Ohio State University
1958 Neil Avenue, Columbus, OH-43202, USA

Abstract

Data capture is the most expensive part of establishing a Geographic Information System (GIS). A Mobile Mapping System the <GPSVan>, integrating a stereo vision system, Global Positioning System (GPS) and Inertial System, has been developed to quickly and accurately collect data. The vision system takes stereo images, while the GPS and inertial systems provide the position and rotation of the vehicle. Any object which appears on an image pair can be located in a global coordinate system. In this paper, we present the mathematical model of camera calibration, the correction of the offsets between different components, as well as the analysis of the positioning accuracy of the GPSVan.

KEY WORDS: Accuracy analysis, Camera calibration, Data capture, System integration, GPS.

1. Introduction

The creation of a geographic information system (GIS) requires an enormous amount of digital information. To date, most land-related databases still rely on existing line maps which are manually digitized. In order to collect digital data faster and more accurately, a mobile mapping system <GPSVan> was developed at The Center for Mapping of the Ohio State University (Bossler, 1991).

This system consists of three major components:

* the control module
* the positioning module
* the image module

Control Module

A PC-compatible computer controls the data collection module through a real-time, multi-tasking operating system. During field operation, a color-touch screen displays the sensor status. The touch screen is also used to note and record attribute information on objects such as bridges and street signs.

Positioning Module

A Navstar Global Positioning System (GPS) receiver determines the global location of the GPSVan. For standard road mapping applications, the GPSVan uses single-frequency GPS receivers in differential mode. Depending on the type of GPS receiver and the operational techniques used, position accuracy can range from three meters to one centimeter.

Because obstructions such as bridges, trees, tunnels or high-rise buildings can interrupt satellite signals, a dead reckoning (DR) system supplements the GPS receiver by recording the distance traveled and the direction of the GPSVan. The DR system consists of a vertical and directional gyroscope and wheel counters and accurately maps the position of the GPSVan in the absence of satellite positions for distances longer than a mile. In addition, it provides the rotation angles of the GPSVan.

Image Module

The image module consists of a stereo vision system, Super VHS color video cameras, and a data collection software package.

The stereo vision system consists of two, fully digital, high resolution CCD cameras (Kodak DCS) with a sensor of 1280x1024 pixels. It acquires image pairs covering the road environment in front or in back of the GPSVan. The stereo vision system is an accurate tool for positioning objects in three-dimensional space. By applying photogrammetric triangulation techniques, any point that appears in both images can be reconstructed in space.

The prerequisite for precise positioning is the system calibration which determines the parameters that define the camera geometry, the relative location and attitude of the camera pair, as well as the relationships between the stereo vision and positioning system.

Image pairs are taken sequentially while the mobile mapping system drives at normal speeds. Through time, every image pair is tagged with the position and rotation of the van in a global coordinate system. Any information extracted from the image pairs is immediately available in the unique global coordinate system by the following two-step transformation:

* Calculation of a local coordinate from left and right image coordinates
* Transformation of a local coordinate to a global coordinate

Fast, accurate acquisition of digital data is the major purpose of the mobile mapping system. In the following, we present the calibration of the GPSVan and the analysis of the positioning accuracy of the system.

2. System Calibration

The calibration of the GPSVan consists of camera calibration, relative orientation and rotation offset determination. The camera calibration is performed by analytical methods which include: capturing the images with different position and view angles of known control points from the test field, measuring all image coordinates, and performing computations to obtain camera parameters. The relative orientation and rotation offset are determined using constraints.

2.1 Camera Calibration:

The calibration itself is done by the well-known bundle adjustment method. The camera parameters which are treated as unknowns are calculated using known control points based on the collinearity equation. The camera parameters are defined by the focal length (c), the principal point (xp, yp) and the lens distortion. We use six distortion parameters, specifically, two for radial distortion, two for decentering distortion and two for affine transformation. The lens distortion is defined by:

$$
\begin{aligned}
dx &= x(r^2-1)a_1 + x(r^4-1)a_2 + \\
&\quad (r^2-2x^2)a_3 + 2xya_4 + xa_5 + ya_6 \qquad (1)\\
dy &= y(r^2-1)a_1 + y(r^4-1)a_2 + \\
&\quad 2xya_3 + (r^2-2y^2)a_4 - ya_5
\end{aligned}
$$

where

a_1, a_2	Radial distortion
a_3, a_4	Decentering distortion
a_5, a_6	Affine parameters
r	Distance to the principal point

When we process the calibration data, all images are defined in a common coordinate system by their positions (X, Y, Z) and three rotation angles. Different rotation systems, e.g. $(\varphi, \omega, \kappa)$, $(\omega, \varphi, \kappa)$ etc. (Kraus, 1992), are available, and we chose the $(\omega, \varphi, \kappa)$ system which is the most popular one in photogrammetry. Collinearity equations are defined by :

$$
x = x_p + dx - c\frac{Nx}{Nz} \qquad y = y_p + dy - c\frac{Ny}{Nz} \qquad (2)
$$

with :

$$
\begin{aligned}
Nx &= r_{11}(X-X_0) + r_{21}(Y-Y_0) + r_{31}(Z-Z_0)\\
Ny &= r_{12}(X-X_0) + r_{22}(Y-Y_0) + r_{32}(Z-Z_0)\\
Nz &= r_{13}(X-X_0) + r_{23}(Y-Y_0) + r_{33}(Z-Z_0)
\end{aligned}
$$

c	focal length,
x_p, y_p	Coordinate of principal point
x,y	Image coordinate
X,Y,Z	Coordinate of an object point(targets)
X_0, Y_0, Z_0	Perspective center of the camera
$r_{11}, \ldots, r_{33}$	Elements of rotation matrix
dx,dy	Correction terms (additional parameters) defined by (2).

All camera parameters are treated as unknowns in equation(2). With a least squares solution, the camera parameters, the position and rotation of every image can be computed using known control points.

2.2. Relative Orientation

In our stereo vision system, two cameras are mounted on a stationary platform. This means that the relative relationship between the two cameras is constant. There are two ways to determine the relative orientation:

a) Stereo camera calibration with relative orientation constraints. This is done by introducing six constraints in the camera calibration procedure to keep the relative orientation and the scale of all stereo pairs constant.

b) An individual relative orientation based on the co-planarity condition.

2.2.1 Relative Orientation Constraints

The relative orientation constraints are introduced to the camera calibration procedure[2.1]. The analytical formulation of this problem is based on the relative rotation matrix ΔR and the relative base vector B between left and right images of any stereo pair. The relative rotation matrix is computed as:

$$\Delta R = R_L^T R_R = \begin{pmatrix} \Delta r_{11} & \Delta r_{12} & \Delta r_{13} \\ \Delta r_{21} & \Delta r_{22} & \Delta r_{23} \\ \Delta r_{31} & \Delta r_{32} & \Delta r_{33} \end{pmatrix} \quad (3)$$

where :

R_L	Rotation matrix of left image
R_R	Rotation matrix of right image

There are only three independent variables in the relative rotation matrix ΔR. We use three rotation angles $(\Delta\omega, \Delta\varphi, \Delta\kappa)$. A constant ΔR matrix is achieved by keeping the relative rotation angles $(\Delta\omega, \Delta\varphi, \Delta\kappa)$ constant. Only the elements $\Delta r_{11}, \Delta r_{12}, \Delta r_{13}, \Delta r_{23}, \Delta r_{33}$ are needed for further computation. Their functional relationship with the rotation angles is given by :

$$\begin{aligned} \Delta r_{13} &= \sin\Delta\varphi \\ -\Delta r_{12} / \Delta r_{11} &= \tan\Delta\omega \\ -\Delta r_{23} / \Delta r_{33} &= \tan\Delta\kappa \end{aligned} \quad (4)$$

These three angles are equaled for every image pair. Assuming that image pair (i) has a relative rotation matrix $\Delta R^{(i)}$, and image pair (k) has the corresponding matrix $\Delta R^{(k)}$, they must satisfy the following equations :

$$\begin{aligned} \Delta r_{13}^{(i)} &= \Delta r_{13}^{(k)} \\ \Delta r_{23}^{(k)} \Delta r_{33}^{(i)} &= \Delta r_{23}^{(i)} \Delta r_{33}^{(k)} \\ \Delta r_{12}^{(k)} \Delta r_{11}^{(i)} &= \Delta r_{12}^{(i)} \Delta r_{11}^{(k)} \end{aligned} \quad (5)$$

The relative base vector B is the base vector defined in the left image coordinate system:

$$B = \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} = R_L^T \bullet \begin{pmatrix} X_R - X_L \\ Y_R - Y_L \\ Z_R - Z_L \end{pmatrix} \quad (6)$$

With

R_L	Rotation matrix of left image
(X_L, Y_L, Z_L)	Position of left image
(X_R, Y_R, Z_R)	Position of right image

The base vectors are also equaled for every image pair :

$$\begin{aligned} b_X^{(i)} &= b_X^{(k)} \\ b_Y^{(i)} &= b_Y^{(k)} \\ b_Z^{(i)} &= b_Z^{(k)} \end{aligned} \quad (7)$$

After linearization, equations (5) and (7) are used as constraints in the bundle adjustment to improve the accuracy of the camera parameters (He, Novak, 1993).

2.2.2 Co-planarity Condition

The co-planarity condition means that the two conjugate image points and the two perspective centers are in one plane. It is defined by :

$$\begin{vmatrix} b_x & b_y & b_z \\ u & v & w \\ u' & v' & w' \end{vmatrix} = 0 \quad (8)$$

where (u,v,w) and (u',v',w') are the three-dimensional image coordinates of the left and right images. We selected the base vector (1, by, bz) and three rotations $(\Delta\omega, \Delta\varphi, \Delta\kappa)$ of the right image in the left image coordinate system as the relative orientation parameters. After linearization of this equation, the relative orientation can be computed easily. No control points are necessary for this computation. For the GPSVan, this method is used to check or update the relative orientation parameter when the stereo vision system is reinstalled.

2.3. Rotation Offset Determination

The rotation offset angles are defined as the rotation difference between the stereo vision system and the positioning system. In the GPSVan, they are small angles (< 5 degrees for most cases). The only way to determine them is by an analytical method. The following constraints are used to compute rotation offsets:

* Same points measured in different image pairs have the same X, Y, Z coordinates
* Points with the same elevation measured from a single image pair have the same Z coordinate.

Our method is based on the following coordinate transformation equation of the GPSVan:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{global} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{gps} + R_{vanori} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{van} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{gps} + R_{vanori} \left[R_{offset} R_{sysrot} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{vision} + \begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix} \right] \quad (9)$$

This equation defines how a point can be transferred from the stereo vision system into the global coordinate system. We want to determine the rotation offset between the stereo vision system and the positioning system. These offsets are small angles and can be expressed by $(d\alpha, d\xi, d\kappa)$. The first derivative of rotation angles is used in the rotation matrix:

$$R_{offset} = \begin{pmatrix} 1 & -d\alpha & d\kappa \\ d\alpha & 1 & -d\xi \\ -d\kappa & d\xi & 1 \end{pmatrix} \quad (10)$$

Let :

$$\begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{gps} + R_{vanori} \left[R_{sysrot} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{vision} + \begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix} \right] \quad (11)$$

$$\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = R_{sysrot} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{vision} \quad (12)$$

then

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{global} = \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + R_{vanori} R_{offset} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + R_{vanori} \begin{pmatrix} -y_0 & 0 & z_0 \\ x_0 & -z_0 & 0 \\ 0 & y_0 & -x_0 \end{pmatrix} \begin{pmatrix} d\alpha \\ d\xi \\ d\kappa \end{pmatrix} \quad (13)$$

This is a linear equation with three unknowns and it is used to form the constraints to derive the rotation offsets.

A point measured from image pair i and j forms three constraints:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}^{i}_{global} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}^{j}_{global} \quad (14)$$

Constraints (14) determine parameter $d\alpha, d\xi$.

For the same elevation points m and n, their Z coordinates are equaled by :

$$Z^{m}_{global} = Z^{n}_{global} \quad (15)$$

Constraint (14) determines $d\xi, d\kappa$.

Collecting all the equations in (14) and (15), the rotation offsets are computed by a least squares solution.

3. Accuracy Evaluation

3.1 Positioning Accuracy of the Stereo Vision System

The algorithms described above were implemented on our post-processing workstation. To evaluate the positioning accuracy, calibrations with different camera parameters and constraints were conducted. All tests were based on two Kodak DCS digital CCD cameras, which were part of the vision system on the GPSVan. They have a resolution of 1280(H) X

1024(V) pixels with a pixel size of 16 micron and a focal length of about 20 mm.

The calibration of the stereo vision system was done using a targeted test field. The stereo pairs were taken at different distances and view angles. Thirteen control points and four check points were used.

3.1.1 Calibration with and without the Relative Orientation Constraints

This test was conducted to analyze the relative orientation constraints. Eight image pairs were used for this purpose. The adjustment was calculated twice; with and without the relative orientation constraints. The relative orientation parameters were derived for every image pair from the exterior orientation parameters obtained by the bundle solution. Tab. 1 and Tab. 2 show the calibration results with and without the relative orientation constraints. Without the constraints, the relative orientation parameters are different from one image pair to another and they are inaccurate.

The base length between the stereo cameras obtained with the relative orientation constraints is exactly the same as the length measured by tape (1.827m ± 1 mm). It is obvious that the calibration with the relative orientation constraints gives the stable parameters. The relative orientation constraints improved the results especially when the geometry of the blocks and the distribution of control points are not good.

3.1.2 Calibration with and without Additional Parameters

In this case, the calibration of the stereo vision system was done using the eight digital stereo pairs of our test field. The bundle triangulation with the relative orientation constraints was computed twice, with and without additional camera parameters. The principal point and the focal length of both cameras were always treated as unknowns. Seventeen points are used in the adjustment, where thirteen were control points and four were used as check points to compare the accuracy. Results are presented in Table 3.

The additional parameters improved the accuracy by a factor of two. The radial distortion parameter k1 is the most significant parameter.

3.1.3 Position Accuracy of the Stereo Vision System

We used the computed orientation parameters in an intersection program to determine the object coordinates from the image coordinate measurements. This corresponds to the positioning of points with the stereo-vision system of the GPSVan, independent of object space control. Again, the coordinate of the targets of the test field were used for comparison. The results are displayed in Tab. 4 showing the RMS errors positioned from different distances.

It is fair to say that the positioning accuracy with two Kodak DCS cameras are in the 10 cm level for objects closer than 30 m in front of the van.

3.2 Global Positioning Accuracy

To test overall accuracy of our mobile mapping system, we measured four control points with a global accuracy of ±1cm. To check the positioning accuracy, all control points were located by the mobile mapping system. Table 5 shows the difference from points positioned by the GPSVan's stereo vision system and their true values. The distances of points to the GPSVan are shown. All points are in the State Plane coordinate system (Ohio south, zone code 3402).

For this test, the Turbo Rogues GPS receivers were installed on the GPSVan. The rotation angles of the GPSVan were taken from the combined adjustment of GPS and inertial system. The offset between the stereo vision system and positioning systems was calibrated using the method discussed in section[2.3].

Image Pair	Bx [m]	B_Y [m]	Bz [m]	$\Delta\omega$	$\Delta\varphi$	$\Delta\kappa$
All	1.826	-0.008	-0.060	-0.04297	0.05933	0.02015

Tab. 1 Relative orientation parameters obtained by adding the relative orientation constraints in a bundle solution

Image Pair	Bx [m]	B_y [m]	Bz [m]	$\Delta\omega$	$\Delta\varphi$	$\Delta\kappa$
200	2.165	-0.028	-0.442	-0.00836	0.05255	0.01837
201	1.647	-0.098	-0.236	-0.00609	0.03777	0.01800
202	1.853	0.033	-0.136	-0.01108	0.04597	0.01712
203	1.848	0.026	-0.135	-0.01378	0.04619	0.01799
204	1.841	0.017	-0.107	-0.01159	0.04594	0.01777
206	1.865	0.011	-0.159	-0.01069	0.04648	0.01866
208	1.861	0.083	-0.145	-0.01607	0.04674	0.01744
209	1.881	0.043	-0.109	-0.01423	0.05041	0.01844

Tab. 2 The relative orientation parameters are different without the relative orientation constraints

	Number of Check Points	S_x mm	S_y mm	S_z mm	σ mm
no ADP	4	1.7	3.2	6.2	6.8
ADP	4	1.6	1.0	3.1	3.7

Tab. 3 Comparison of calibrations with and without additional parameters (ADP)

Object Distance	No. of Points	S_x [m]	S_y [m]	S_z [m]	σ [m]
30m	5	0.029	0.010	0.115	0.119
17m	9	0.010	0.014	0.068	0.070
9m	8	0.005	0.009	0.023	0.025

Tab. 4 Positioning accuracy of the stereo vision system where the object distance corresponds to the Z coordinate

Points	Distance to GPSVan	ΔX [m]	ΔY [m]	ΔZ [m]
A	26.5	-0.125	0.109	-0.160
B	10.6	-0.063	-0.406	-0.275
C	19.7	0.000	-0.109	-0.102
D	20.0	-0.187	-0.406	0.037
σ		0.117	0.297	0.127

Tab. 5 Global positioning accuracy of the GPSVan and stereo vision system

4. Conclusions

In this paper, we have presented all of the algorithmic aspects of GPSVan calibration. The positioning accuracy of the GPSVan is better than one foot in a global coordinate frame. This meets the requirements of most engineering applications. For example, we have finished projects for railroads, airports and utility applications. All projects have been completed successfully and much faster than using conventional surveying techniques. In the future, we plan to use dual frequency receivers and an inertial system to improve the accuracy to better than 10 cm.

5. References

Bossler J.,Goad C. ,John P., Novak, K., 1991. "GPS and GIS Map the Nation's Highways," GeoInfo System Magazine, March issue, pp. 26-37.

Brown D.C., 1976: "The Bundle Adjustment - Progress and Prospects," Invited Paper XIII the Congress of ISP, Commission III, Helsinki.

Novak K. 1991. "The Ohio Sate University Highway Mapping System: The Stereo Vision System Component." Proceedings of the Institute of Navigation Conference, Williamsburg, VA, pp. 121-124.

He, G.P., Novak, K. Feng, W.H, 1993: "Stereo Camera Calibration with Relative Orientation Constraints," in Videometric, Sabry F. EL-Hakim, Editor, Proc. SPIE 1820, pp 2-8.

Zhang, Zuxun, 1990: "Digital Photogrammetry," published by the Wuhan Technical University of Surveying and Mapping, Wuhan, PR China.

Kraus, K. 1992. "Photogrammetry," Dümmler Verlag. Bonn.

Wang, Z.Z. 1990. "Principles of Photogrammetry (with Remote Sensing)," Publishing House of Surveying and Mapping, Beijing.

PRÉCISION DE POSITIONNEMENT OBTENUE AU MOYEN DU GPSVAN

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif

ADVANCED 3D VISUALIZATION TECHNIQUES

COLIN WARE

Faculty of Computer Science, P.O. Box 4400
University of New Brunswick, Fredericton, N.B. E3B 5A3
email: cware@UNB.ca

ABSTRACT

Highly interactive 3D graphics is on the threshold of becoming affordable and this opens the door to a whole new range of applications. However, while a well defined set of interaction techniques exists for the 2D screen environment, the 3D user interface is evolving rapidly. This paper focuses on the user interface to 3D graphics environments. Novel input devices, 3D widgets, Fly-By techniques, Stereo displays, and Fish Tank virtual reality are discussed within the framework of J.J. Gibson's theory of affordances.

L'interface utilisateur pour la visualisation en 3D

Le graphisme en 3D hautement interactif est sur le point de devenir abordable et cela ouvre la porte vers toute une panoplie de nouvelles applications. Néanmoins, alors qu'il existe un ensemble bien défini de techniques pour l'environnement de l'écran en 2D, l'interface utilisateur en 3Dévolue rapidement. Cet expose va se concentrer sur l'interface utilisateur pour un environnement graphique en 3D. Des nouveaux périphériques d'entrées, les widgets en 3D, les techniques de survol, la représentation en stéreo, la réalite virtuelle vont être discutés à travers la notion théorique de J.J. Gibson, 'affordances'.

KEY WORDS: 3D views, Geographical Information System, Visual data display.

1. INTRODUCTION

The great contribution of Geographical Information Systems can be seen as an advance from one dimensional information, usually tables of scalar coordinates, to two dimensional information - a map linked to a data base. The methods for manipulating maps are well established and they consist for the most part of scale and translate operations for positioning, as well as point, line and area specification for data input. A third dimension is added to a limited extent in the form of overlays, providing layers of information. However, overlays are a convention for representation, in general the layer does not denote scalar height information.

This may all be about to change as high performance 3D graphics systems become commonplace. There are increasing numbers of people experimenting with 3D representations of data and it is becoming clear that this is useful and practical for many applications. For example, understanding the surface topography in highway planning, appreciating sight lines in urban planning or forestry, and representing ore bodies in mining applications are all areas where 3D visualization can help. In general 3D visualization is useful where the structures to be understood are physically displaced in the vertical direction and this displacement is critical to the application. Three dimensional representations will become even more important as GISs evolve into GCADS (Geographical Computer Aided Design Systems).

To realize the potential of 3D GIS a number of problems must be solved. In particular, there are new data visualization and manipulation techniques to be invented (after all true 3D output devices are still many years away). While 2D GIS systems have borrowed heavily from the paper maps that preceded them, and which are often the most useful product. 3D GIS has no such antecedents, unless it is the model of the landscape architect. There are other problems related to the need for a the database to support true 3D visualization and to retain volumetric information, and we need to support true 3D queries into that database. These database issues are beyond the scope of this

paper which will address the user interface and visualization issues.

Before continuing, a small theoretical digression is in order. One of the most useful concepts in defining quality in user interfaces is J.J. Gibson's theory of affordances (Gibson, 1966). This is theory of perception which states that we do not perceive patterns of light and shade, or lines and edges, or motion flow. Instead what we have evolved to perceive are 'affordances' which may be described as possibilities for action, or use. This we perceive surfaces as having the potential for walking or sitting, we perceive objects as potential tools or potential food and certain complex environments as holding potential danger. The theory embeds perception in action, and as such it makes sense of a good user interface as one in which the user perceives the right set of affordances. Through it, with minimal instruction the user can perceive the affordances of the computer system he or she is expected to use. The system should also afford the easy execution of the tasks for which it was designed.

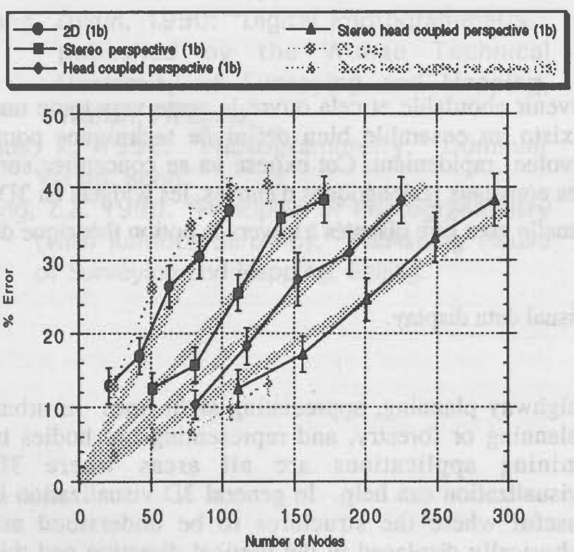


Figure 1. The results of a study of path tracing in an information network. Using a stereo, head coupled perspective view, as shown in figure 2. resulted in three times as many nodes being understood at the same error rate.

We have recently obtained hard evidence that even for visualizing abstract information networks, where understanding the connectivity is important, visualizing in 3D is important (see Figure 1). However, it is not the fact that it is a perspective view that helps, but rather the enhanced space perception that comes from stereo viewing (which increases the size of the network that can be understood by 60%) and even more from motion parallax of the data (which increases the size by 120%). If motion parallax is combined with stereo viewing we find that three times the network size can be understood for a constant error rate.

2. METHODS AND METAPHORS FOR VIEWPOINT NAVIGATION

Viewpoint placement, 3D scene exploration and virtual camera control are all aspects of the same problem in computer graphics, namely how to move the viewpoint in a virtual 3D scene. The kinds of task where this is important are molecular modeling (Surles, 1992), walkthroughs of architectural simulations (Brooks, 1986), camera control in animation systems, and flights over digital terrain maps representing subsea or remote sensing data (Stewart, 1991) as well as numerous CAD and advanced GIS applications.

For a number of years we have been studying a six degree-of-freedom variant of the common mouse input device. We call it a Bat because a bat is like a mouse that flies (or *fledermaus* in German). The device senses both position (x,y,z) and orientation (azimuth, elevation and roll) information. In some studies we showed how this device could be used for object placement (Ware, 1990). However, more recently we have concentrated on using the Bat in ways that allow us to explore different methods and metaphors for virtual camera control.

Often methods for viewpoint control are based on metaphors which help the user to get a conceptual grasp of the way the system will behave. Thus if the user is told that he or she is flying through the data, it is quite different than telling the user that the data is on a turntable which can be rotated. Most of the remainder of this paper is organized as a survey of different virtual camera control methods, both as employed in my research laboratory and by others.

1) Eyeball in Hand Metaphor and Camera controllers. The phrase "Eyeball In hand" describes a Metaphor in which the user directly manipulates the viewpoint as if it were held in his or her hand. The metaphor requires that the user imagine a model of the virtual environment somewhere in the vicinity of the monitor. The eyeball (a spatial positioning device) is placed at the desired viewpoint and the scene from this viewpoint is displayed on the monitor. Cognitive affordance problems arise from the difficulty some subjects have of imagining the model. Ware and Osborne, (1990) found large individual differences in this respect. Also, if the eyeball is pointed away from the screen the correspondence between hand motion and the image motion is confusing. Physical affordances are restricted by the physical limitation of the device space - it can be awkward or impossible to place the "eyeball" in certain positions.

There is a non direct-manipulation variation on this theme which allows for complex camera commands of the kind a director might give to the cameraman. Recent work by Gleicher and Witkin (1992) explores the use of high level commands to give the user control over the viewpoint in a manner which is similar to directing a cameraman to make a movie shot . Thus the director might request a cameraman to track a certain moving object, pan from object A to object B, or zoom in to a close-up. An even more ambitious approach is to have an autonomous smart camera which positions the viewpoint automatically to allow the user to concentrate on some other task (Phillips, et al, 1992).

2) World in Hand and Mechanical Metaphors.

In the "World in Hand Metaphor" the viewpoint is changed by moving the object. Thus, to look at a displayed scene from the right hand side, the scene is rotated clockwise (as viewed from the top), like an object on a turntable. If a 6DF input device, such as the Polhemus™, or a SpaceBall™, is available then rotations and translations can be carried out simultaneously. As a method for changing the viewpoint the world in hand metaphor works well for single, reasonably compact objects. However, when the environment to be explored is a landscape or enclosed interior space, picking it up and moving it does not seem natural (it has poor cognitive affordances). There also exists a problem selecting the center of rotation. Especially when moving through an interior the metaphor clashes with the user's perception of being enclosed and the linkage of scene motion to hand motion is incongruous and difficult to grasp (Ware and Osborne, 1990).

The World in hand metaphor is almost the opposite of the eyeball in hand metaphor; instead of moving his or her viewpoint the user imagines moving the object. Useful variants on the metaphor are such devices as the virtual turntable (Evans, et al, 1981), stirrer, or virtual sphere (Chen, et al, 1988) which transform the input from a mouse or digitizing tablet to give the feeling of direct manipulation of a graphical object. These devices tend to be easy to learn but not very flexible. Often they must be carefully customized to provided the range of movements required by the application. We are using a virtual turntable in an interface to a DEM visualization system for satellite imagery.

3) Functions and Smart functions.

It is a common practice to control the viewpoint using common graphics library functions such as scale, translate and rotate. These functions may be controlled directly by the mouse or indirectly via sliders. We include this kind of interface for completeness although it does not embody the use of a consistent metaphor.

A much more interesting use of (non metaphorical) functions is the point of focus zoom developed by MacKinlay, et al (1990). The implemented a zoom which corresponds to moving the viewpoint to the surface, halving the distance for each unit time. The also evaluated some complex compound viewpoint movements, for example, to zoom in on a particular surface point and at the same time rotate so as to place the surface at right angles to the viewing direction.

4) Head Coupling and the Virtual Reality (VR) metaphor

The VR metaphor involves coupling the perspective image to the user's head position so that, for example, to look at the far side of an object the user must walk around the object (Sutherland, 1968). This method uses the interface of everyday life but its affordances are highly restrictive, it only allows viewpoint manipulation within the range of head movements. To allow for greater flexibility it must be combined with other methods. It most naturally fits with the use of mechanical widgets (such as a virtual turntable) because they can be implemented as objects in the virtual world (Connor, et al. 1992). It is also possible to implement a localized form of VR using a conventional workstation an coupling the viewpoint to the measured eye position (Ware, et al, 1993) as shown in Figure 2.

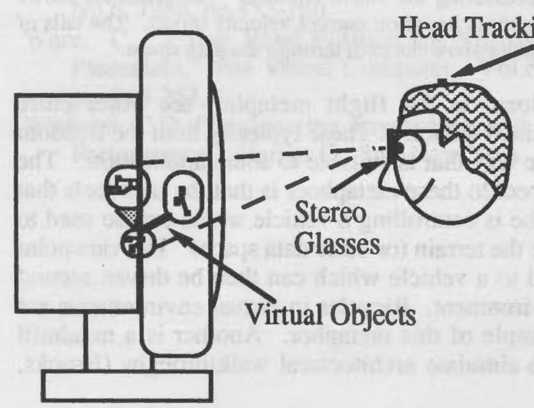


Figure 2. By coupling the perspective viewpoints for the two eyes to the measured head position of the observer and using stereo glasses it is possible to create a highly realistic localized virtual image.

5) Flying and virtual vehicle control

Flight simulators have the rather inconvenient affordances of flying vehicles, problems such as stalling when the velocity is low are hindrances to arbitrary control of viewpoints. We have done extensive work towards making an easy-to-use velocity control interface for exploring 3D graphical environments. This which makes no attempt to model real flight dynamics but is designed to give a great flexibility in the control of viewpoint movements through environments which mostly consist of scientific data representations (Ware and Osborne, 1990). Predictive feedback can greatly enhance the user's sense of control (Chapman and Ware, 1990).The form of predictor that we have developed is illustrated in Figure 3. We use this system in making movies to convey information to others.

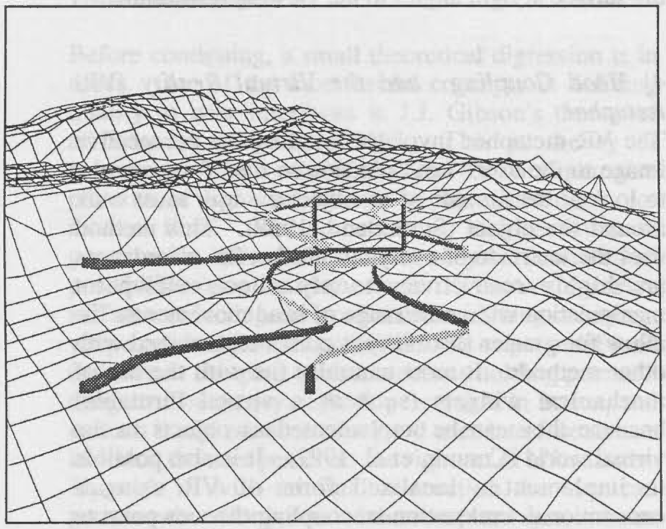


Figure 3. The predictor is seen in use over a Digital Terrain Map representing the North Atlantic. The predictor shows future position based on current velocity inputs. The tails of the predictor show the path through the data space.

Variations on the flight metaphor are other more constrained vehicles. These typically limit the freedom in some way that is suitable to some application. The key aspect to these metaphors is that the user feels that he or she is controlling a vehicle which can be used to explore the terrain (or other data space). The viewpoint is fixed to a vehicle which can then be driven around the environment. Bicycles in virtual environments are an example of this metaphor. Another is a treadmill used to simulate architectural walkthroughs (Brooks, 1986).

3. DATA MANIPULATION IN 3D

The problems of data manipulation and viewpoint manipulation both involve six degree-of-freedom task performance. It takes three numbers to position an object in space and three more to specify its orientation. There are many solutions to the problems of 3D object placement, including virtual turntables (Evans et al, 1981), or virtual trackballs (Chen et al, 1988) whereby the mouse is used to control a set of higher level widgets which control rotations and translations. We have had considerable success using the Bat, described above for free object placement. However, in many GIS and CAD operations we find that it is better to used constrained placement. Often the user wishes to only manipulate one degree of freedom at a time, a rotation about a particular axis, or a single axis translation. Therefore, whereas we use the Bat extensively for viewpoint manipulation, we generally use a conventional mouse for object placement in the applications we are currently working on.

4. THE FLEDERMAUS SYSTEM

We (Mark Paton and myself) are currently building a prototype 3D Geographical Visualization System at the University of New Brunswick that incorporates many of the advanced interaction techniques talked about in this paper (it is called Fledermaus). Our primary use of the flying technique is in a data visualization and editing system for oceanographic research. This enables users to fly over a digital terrain map represented in thinned wire frame form, The thinning is necessary because it is impossible to render the entire surface at the highest resolution and at a reasonable upgrade rate. As soon as the user stops flying, the surface fills in two steps: a medium resolution step and the highest resolution step. The idea is to keep adding information to the user and to fill in the otherwise unacceptable delay required to compute a fully rendered high resolution image (on our current IRIS Crimson VGX™ it takes 16 seconds to render a 2000x2000 digital elevation model). In order to make a high quality video of a flyby it is only necessary to hold down a different button on the Bat and from that time forward the sequence of viewpoints is recorded while flying. These viewpoints are later used to generate high resolution frames which are captured one at a time on our Panasonic LQ4000 laser disc recorder. The videos we have made in this way (using previous versions of the system) have appeared in the Canada pavilion at the recent Seville World's fair, on Canadian and German National Television as well as at numerous scientific and technical conferences.

We are currently adding Georeferencing, to the visualization system and it now has support for multiple objects. This enables us to, for example, compare two different DEMs of the same surface, by co-registering them and moving them up and down with respect to each other. We have also recently added support for closed polygon meshes with the goal of visualizing schools of fish in the water column. We are currently investigating various techniques for stereo visualization and by the time this paper is presented should have stereo movie capabilites.

ACKNOWLEDGMENTS

I gratefully acknowledge the support and assistance of Dr. Larry Mayer who has been instrumental in encouraging the practical application of laboratory ideas. The research described here has been primarily funded through NSERC grants.

REFERENCES

Brooks, F.P. "Walkthrough - A Dynamic Graphics System for Simulating Virtual Buildings," Proc 1986 Workshop on Interactive 3D Graphics, F.

Crow and S.M. Pizar, eds, ACM, New York, pp.9-21.
Chapman, D. and Ware, C. Manipulating the Future: Predictor Based Feedback for Velocity Control in Virtual Environment Navigation. Special Issue of Computer Graphics. Proceedings of 1992 Symposium on Interactive 3D Graphics, Cambridge Mass, 63-66.
Chen, M. Mountford, S.J. and Sellen, A. "A study of Interactive 3D rotation using 2D control devices," Computer Graphics (Proc. Siggraph), Vol.22, No.4, Aug.1988, pp. 121-129.
Connor, D.B. Snibbe, S.S. Herndon, K.P. Robbins, D.C. Zeleznik, R.C. van Dam, A." Three-Dimensional Widgets," 1992 Symposium on Interactive 3D graphics April, 1992, Special Issue of Computer Graphics, pp. 183-188.
Evans, K.B. Tanner, P.P. and Wein, M. "Tablet Based Valuators that provide one, two, or three degrees of freedom," Computer Graphics (Proc. Siggraph), Vol.15, No.3, Aug. 1981, pp. 91-97.
Gibson, J.J. (1966) The Senses Considered as Perceptual Systems. Houghton Mifflin: Boston.
Gleicher, M. and Witkin, A. "Through-the-Lens Camera Control, Computer Graphics (Proc. Siggraph), Vol. 26, No. 2, Aug. 1992. pp. 331-340.
Mackinlay, J.D. Card, S.K. and Robertson, G.G. "Rapid Controlled Movement Through a Virtual 3D Workspace," Computer Graphics (Proc. Siggraph), Vol.24, No.3, Aug. 1990, 171-176.
Phillips, C.B., Badler, N.I. and Granieri, J. "Automatic viewing control for 3D direct manipulation," 1992 Symposium on Interactive 3D Grapics. Special issue of Computer Graphics, March 1992, pp.71-74.
Stewart, W.K. "Multisensor Visualization for Underwater Achaeology," IEEE Computer Graphics and Applications, Vol.11, No.2, Mar. 1991, pp. 13-18.
Surles, M.C. "An Algorithm with Linear Complexity for Interactive Physically-based Modelling of Large Proteins," Computer Graphics (Proc. Siggraph), Vol.26, No.2, July 1992, pp. 221-230.
Sutherland, I.E. "Head Mounted Three Dimensional Display. Proc of the Fall Joint Computer Conference," Vol.33, 1968,757-764.
Ware, C. Arthur, K. and Booth, K.S. (1993) Fishtank Virtual Reality. INTERCHI'93 Technical Paper. Proceedings 37-42.
Ware C. and Osborne, S. "Exploration and virtual camera control in virtual three dimensional environments," Proceedings of the 1990 Symposium on Interactive 3D Graphics (Snowbird, Utah). In Computer Graphics Vol.24, No. 2, March 1990, pp.175-183.
Ware C. and Jessome, D.R. "Using the Bat: A Six Dimensional Mouse for Object Placement," IEEE Computer Graphics and Applications, Vol. 8, No.6, 1988, pp.155-160.
Ware, C. "Using Hand Position for Virtual Object Placement," The Visual Computer, Vol.6, 1990, pp. 245-253.
Wickens, C.D. "Engineering Psychology and Human Performance," Harper Collins: New York. 1992

GEOGRAPHIC INFORMATION: THE WAY AHEAD
"Ostendamus Viam - We Show The Way"

David G. McKellar

Department of National Defence
Directorate of Geographic Operations
National Defence Headquarters
Ottawa, Ontario, Canada K1A 0K2
Telephone: (613) 995-4239
Fax: (613) 996-3328

Abstract

In recent years significant progress has been made in standardization and interoperability with respect to geospatial data. Along with this progress has come a corresponding increase in requirements from all military environments. There is now a biennial effort underway within NATO to solicit user requirements for geospatial data.

Due to the success in the acceptance of the DIGEST standard at national, regional and international levels, it is now possible to develop a set of common software tools (DIGEST View) based on standard open system Application Programmers Interfaces (APIs). This tool box of public domain source code will be available to users and integrators to develop their applications.

In order to ensure that geospatial data is available when required, the allied military geographic agencies have undertaken a massive production programme known as Vector Smart Map (VMap) at various levels of resolution to support, for example, a stated NATO requirement for a Background Display Dataset (BDD).

An integral aspect of this standardization and associated geospatial data production is a healthy research and development effort. Ongoing work is required in the area of data fusion, standards, interoperability and simulation.

Success will only be achieved through ongoing efforts to harmonize standards efforts for geospatial data, and by forming strategic partnerships with industry and academia.

Information géographique - Perspectives «Ostendamus viam» - Nous indiquons la voie

RÉSUMÉ

Au cours des dernières années, d'importants progrès ont été réalisés en ce qui a trait à la normalisation et à l'interopérabilité en matière d'information géospatiale. Ces progrès se sont faits parallèlement à une demande croissante d'information de la part des utilisateurs militaires. On tient maintenant chaque semestre, au sein de l'OTAN, une campagne visant à inciter les utilisateurs chargés de missions ou de fonctions particulières à faire part de leurs besoins en matière d'information géospatiale.

L'acceptation rapide et générale de la norme DIGEST, aux niveaux régional, national et international, a permis la mise au point de logiciels-outils standard (DIGESTVIEW), élaborés à partir d'une interface normalisée de programmeurs d'application (API). Cette boîte à outils de codes sources sera mise à la disposition des utilisateurs et des intégrateurs pour leur permettre d'élaborer leurs applications.

Dans le but d'assurer que l'information géospatiale soit accessible sur demande, les organismes d'information géographique des forces alliées ont entrepris un programme de production massive appelé Vector Smart Map (VMAP), à différents degrés de définition, afin de répondre par exemple à une spécification de l'OTAN en vue d'un dispositif d'affichage d'ensembles de données d'arrière-plan.

Tout cet effort de production de données géospatiales et de normalisation a pour conséquence de stimuler la recherche et le développement. Les travaux doivent se poursuivre en ce qui a trait à la fusion des données, aux normes, à l'interopérabilité et à la simulation ou à la modélisation.

Il est essentiel à la réussite de ces travaux qu'on mette tout en oeuvre afin d'harmoniser les normes relatives à l'information géospatiale et à former des alliances stratégiques avec des partenaires des milieux universitaires et de l'industrie.

1. INTRODUCTION

Reality is that the allied military and intelligence communities need solutions today. This is achieved by initially populating data bases through conversion of existing analog products into a digital form. This then must be highly standardized to ensure true interoperability in joint or combined operations. Current users of geospatial data are not normally sophisticated GIS experts and therefore, expect information to be provided in a totally non ambiguous manner. Subsequent editions of these products will be compiled from new centre line data. As producers we are faced always with doing more with less, therefore it is imperative that todays production systems are capable of producing simultaneously both soft and hardcopy products. Even though future systems will be more sophisticated and responsive, there will still be a stated requirement to support todays standard geospatial products. This will remain a challenge for system developers.

2. STRATEGIC DIRECTION

It is imperative that we must recognize and acknowledge enabling technologies such as microelectronics, satellite navigation, telecommunications, computer power and storage, sensors, data basing and networking. Today, and for the next several years, we will continue to provide a number of standard products with specific encapsulations and defined media. These will be used by fielded systems or preprocessed prior to use. The future is accessible online or near online data bases with data available in a standard exchange format which can be visualized using standard tool boxes or utilities. It is expected that these data bases will use either object-oriented or relational Data Base Management System (DBMS) concepts.

3. CLIENT/USER SERVICES

There is a need for improved accuracy, precision and currency, along with responsiveness. An efficient and effective feedback mechanism is essential to allow the user to query an inventory or catalogue, metadata and requirements status. Another aspect which has to date received very little attention, is the real time updates and maintenance issue. This will require a bidirectional process to accept value added data and validate it in near real time. Education and training of the user is equally important, but historically does not receive the priority that it should.

4. STANDARDS AND SPECIFICATIONS

The area which everyone recognizes as critical, yet are reluctant to participate or invest in is standards. Investment in standards is imperative critical for without them we will continue to face the unsatisfactory situation which we all now must live with.

It must be recognized that every standardization effort that is underway at the national, regional or international level can not succeed. Harmonization amongst these efforts is mandatory if we are to reach our shared objective. Convergence to one family of geospatial standards is not impossible. The challenge is to make this objective a reality.

Within Canada, under the auspices of the Canadian General Standards Board (CGSB) Committee on Geomatics (COG), we have created a formal standards process. To date we have developed several standards, the Geomatics Data Sets: Cataloguing Rules and another covering directories of geomatics data bases. With respect to exchange standards we have developed a mechanism to handle both de jure and de facto standards. The COG recently adopted both SAIF, developed by the province of British Columbia, and DIGEST, as two of the Canadian Geomatics Interchange Standards (CGIS). SAIF has been categorized as a general standard while DIGEST is a defined standard. Both of these efforts can coexist by developing registered profiles within the general standard compliant with the defined standard. Other ongoing activities involve establishing standards for a national feature and attribute catalog.

Similarly, other national standards boards have in place comparable efforts. Two examples of regional efforts is that of the European Committee for Standardization (CEN) Technical Committee 207 on Geographic Information and the NATO Military Agency for Standardization (MAS), Interservice Geographic Working Party. At the international level, there are several concurrent activities ongoing. Three such efforts are; the International Hydrographic Organization (IHO) and the S-57 standard, the Digital Geographic Information Working Group (DGIWG) and the DIGEST standard, and the most recent activity being the formation of an ISO Technical Committee on Geomatics, based on a Canadian proposal. It is through the latter where true standardization efforts will take place. Therefore, it is imperative that people invest in geospatial standards now while the opportunity exists to influence the outcome. It is not well known that Canada is the custodian nation for DIGEST both within DGIWG and NATO, and perhaps in the future in ISO.

Most of the efforts today, even the most progressive ones, are too conservative or pessimistic. Discussion is still about common data sharing architectures and how to handle the legacy and propriety data bases that exist today. Perhaps in the military geographic community, because of our increased world wide mandate, we tend to take a more proactive or aggressive view. We feel that it is possible to address the exchange of standard geospatial data between producers, producers and users, and between users. If the standards process is in place, the standards are developed, vast amounts of data are readily available, and the tools to exploit them, then it is realistic to envision a single or at the worst a very small number of standards for geospatial data. A trend is already apparent as leading GIS companies and third party integrators develop robust bidirectional translators, some have even gone as far as to rewrite their GIS engine to be compliant to emerging world standards, such as DIGEST.

One of the greatest areas for confusion is in the area of terms and definitions. If anything will bring a standards process to a halt it is this problem. Within the international Digital Geographic Information Working Group (DGIWG) we have struggled with some very basic concepts with respect to data. First of all we have differentiated between the term data and information. Data is defined as "instances of facts with specific meanings occurring in the real world". On the other hand, information is "a grouping of data in context related to a specific purpose .... a formalized accumulation of data to resolve uncertainty". Neutral data is "data which is not intended to support a specific application unless combined with other information". In other words individual data elements may be neutral. The Feature Attribute Coding Catalog (FACC) is a neutral dataset specification. Annex C of FACC, in DIGEST 1.2, is a first draft of a neutral dataset. For the time being, efforts on further definition and management of neutral datasets has been suspended. The next term used is a DIGEST dataset, which is "a collection of data arranged in a DIGEST compliant format .... normally used for bulk transfer between producers". From there we can define a standard DIGEST dataset, which is "a collection of data which has a specification which pre-defines the content and the DIGEST encapsulation". An example of standard DIGEST dataset is the Background Display Dataset (BDD) which has specified contents for level 0 (>1:1,000,000), level 1 (1:250,000 - 1:1,000,000) and level 2 (1:50,000 - 1:250,000). The last term used is a DIGEST product, which is "a standard DIGEST dataset with a defined media and packaging, stated in a product specification .... usually delivered to an end user". An example would be the Vector Smart Map (VMap) series which meets the BDD standard DIGEST dataset specification.

5. GEOSPATIAL TOOL BOX

One of the key ingredients for the successful adoption of a world standard is the ready availability of a comprehensive geospatial toolbox which adheres to open system architecture. Without a method of visualizing the data users will never buy into standardization or interoperabilty. This allows for maximum reusability of data and reduction in duplication of effort in software development.

DIGEST View is a three year R&D effort within DGIWG, initiated by Canada. This project is built on and in cooperation with a U.S. Defence Mapping Agency effort known as the Mapping, Charting and Geodesy (MC&G) Utility Software Environment (MUSE). The primary objective is to demonstrate multiple system platform interoperability (DOS, MAC, UNIX), and to provide a basic capability to import, display and manipulate matrix, raster, vector and text DIGEST products. Release of the software will be in the public domain including full source code, user manuals and system description manuals. This will allow users and integrators to fully exploit all DIGEST products.

6. INFORMATION HIGHWAY

There is a need to provide a wide range of information services to meet the emerging more sophisticated user and intelligent real time broadband networks. This ranges from a single workstation to netted servers and distributed processing. Ever increasing band widths and massively parallel computing will push geospatial data producers to react quickly to this important infrastructure initiative. Geospatial architectures should be open and easily accessible. Some of the pioneering work in this area has been done by the Open GIS Foundation with the Open GIS Concept.

7. USER ACCESSIBLE DATABASES

Today we organize geospatial data into products, but this may not meet or solve a users systems decisions effectively. For example, any standard dataset consisting of predefined features and attributes organized in structured thematic layers may not be what a user needs to make an informed decision. This is a similar situation to more traditional paper maps and charts where content and presentation were predetermined. A partial solution to this dilemma is used in NATO by following a layered data base approach where a mix and match philosophy creates various data sets for particular applications, such as transportation and logistics or terrain analysis.

Geospatial data bases of the future must be more user accessible. Several significant research and development projects are underway to ensure that users needs are addressed. These projects include extensions to DIGEST to address expanded requirement for raster products both facsimiles of existing maps and charts, but also multispectral and grey scale geocoded satellite images. As the size of geospatial data bases grow with logarithmic rates, so must the means to load and retrieve them efficiently. Work is underway to extend the relational DBMS using HH coding and the new ISO SQL extensions. Both of these efforts will be incorporated as part of a major new release of DIGEST by 1996. The other crucial area of development is in the area of standard open system tool boxes which will provide users and integrators with the tools they need. Research is also ongoing in multiscale - multiproduct data bases which will then allow users the full flexibility they need to carry out their mission. It is envisioned that we will produce in the next several years a variety of DIGEST products such as Vector Smart Map (VMap), Digital Nautical Chart (DNC), and Arc Digital Raster Graphics (ADRG) of various resolutions over the same geospatial extent on the same standard media.

8. CONCLUSION

Without a doubt we are in exciting times where the pace of change is revolutionary, not evolutionary. The challenge remains to cope with today, envision the future, and plan for an orderly and responsive transition.

9. ACKNOWLEDGEMENTS

The author wishes to acknowledge many discussions with colleagues from allied military geographic agencies, and in particular the U.S. Defence Mapping Agency. It is only through close cooperation and collaboration that the way ahead becomes clear.

REFERENCES:

British Columbia Specifications and Guidelines for Geomatics, Reference Series, Volume 1 [1993]: Spatial Achieve and Interchange Format (SAIF): Formal Definition, Release 3.0 Draft.

Digital Geographic Information Working Group [1994]: Digital Geographic Information Exchange Standard (DIGEST) Version 1.2.

Dobson, J.E. [1994]: GIS Integration Requires a New Kind of Parallel Processing. GIS World March 1994, page 28.

European Committee for Standardization (CEN) TC 287 [1993]: Work Programme (N154) and Reference Model Working Draft - Version 2 (N154).

Federal Geographic Data Committee (FGDC) [1994]: The 1994 Plan for the National Spatial Data Infrastructure: Building the Foundation of an Information Based Society. Final Draft.

Keighan, E. [1992]: Managing Spatial Data Within the Framework of the Relational Model. Oracle Corporation Canada, Publication Unknown.

McKellar, D.G. [1989]: GIS Technology in National Defence. First National Symposium on GIS, Ottawa Canada.

McKellar, D.G., Feeley J. [1990] The Digital Chart of the World Project. Second National Symposium on GIS, Ottawa Canada.

McKellar, D.G. [1991]: The Digital Geographic Information Working Group and International Standards. Third International GIS Conference, Ottawa Canada.

McKellar, D.G. [1992]: Management of GIS Standards in an International Context. Fourth International Conference on GIS, Ottawa Canada.

McKellar, D.G. [1992]: The Digital Chart of the World Project: Moving Towards a Global GIS Information Infrastructure. Fourth International Conference on GIS, Ottawa Canada.

McKellar, D.G. [1993]: The International Context. Fifth International Conference on

GIS, Workshop #6 - Issues in Geomatics Standards, Ottawa Canada.

Salge, F. [1993]: Workshop on Strategy for National, Regional, International De Jure Standards in The Field of Geographic Information. Published Minutes, Paris France.

Strand, E.J. [1993]: Digesting the DIGEST Data Exchange Standard. GIS World September 1993, page 22.

The Open GIS Foundation [1993]: The Open Geodata Interoperability Specification (OGIS) Version 1.0, Preliminary Draft

Varma, H. [1992]: Implementation of Topological Relationships Using SQL. Canadian Hydrographic Service, Publication Unknown.

Vezina, G. [1993]: Parallel Spatial Analysis and Interactive Visualization Software. Defence Research Establishment Valcartier, Internal Paper.

A GEOSPATIAL DATA FRAMEWORK FOR THE UNITED STATES

Stephen C. Guptill
U.S. Geological Survey
519 National Center
Reston, VA 22092 USA
phone (703) 648-4520
fax (703) 648-5542
sguptill@usgs.gov

ABSTRACT

The National Performance Review (a study of Federal government operations aimed at making the government more effective, efficient, and responsive, conducted by Vice President Albert Gore during the spring and summer of 1993) has recommended that the Executive Branch develop, in cooperation with State and local governments and the private sector, a coordinated National Spatial Data Infrastructure (NSDI) to support public and private sector applications of geospatial data. In support of this goal, the Federal government shall complete an initial implementation of a national geospatial data framework. The framework will provide a consistent information base for the maximum number of users, against which other data categories (e.g. soils, geology), more detailed content within a data theme, or data from more than one source can be registered. The framework design will provide assurance that all data registered to it will be able to be related to and used with other spatial data within the constraints of its level of accuracy. The framework will aid in locating data to its correct position and provide transparent referencing to other spatial information. It will provide a nationally accepted and used referencing capability that will eliminate recollection, duplicate collection, or incompatible collection of basic building blocks of the NSDI.

UNE STRUCTURE DE DONNÉES GÉOSPATIALES POUR LES ÉTATS-UNIS

Résumé

Traduction non disponible pour cause de livraison tardive du résumé définitif

Delta-X: A Federated Spatial Information Management System

Ekow J. Otoo and Adel Mamhikoff
Centre for Parallel and Distributed Computing
School of Computer Science
Carleton University
Ottawa, Ontario, K1S 5B6

Abstract

Development of applications using Geographic Information Systems often requires access to several different data-sets. These data-sets are heterogeneous with respect to i) the underlying database management system, ii) the spatial and non-spatial data formats, iii) the data storage technology, etc. A considerable volume of the required data is not even maintained with any database technology. The data-sets are stored and delivered as simple files. This may be attributed to the fact that digital cartographic data collection preceded the development of databases technology. As a result, the data accessible to most users is legacy data. Spatial information systems, and in particular GIS technologies, lack the unified integrated access to the multitude of different data types required for application development. The Delta-X federated spatial information management system is an approach to provide interoperability in a network of heterogeneous databases. It establishes interoperability between different relational DBMS's, simple files and object repositories. Using the Delta-X approach, an agency can register a data-set, that is either maintained or not maintained by an underlying database management system. This is made accessible to users through a distributed virtual global schema definition. By this means, the independence and autonomy of control, of the individual data-sets registered in the federation, are still respected, while still maintaining some integrated uniform access with respect to a particular GIS environment.

Keywords: distributed databases, interoperability, relational data model, object repositories.

1 Introduction

Geographic Information Systems (GIS), has become and will grow to be a major application area of database management systems that rides on the information super-highway in the future. Essentially, it is a tool for simulating, modelling, analyzing and visualizing geographic features and geophysical phenomenon of the world around us. Its scope of application is unbounded. Typical application areas include: development of electronic maps, geographic encyclopedia, natural resource exploration and management, emergency dispatching, transportation, utility management, real estate management, city and community development, environmental protection, etc., [2]. GIS's applications impact all spheres of our daily lives.

At the heart of a GIS system is the information base, i.e., a collection of databases that store the geometric definition of spatial features and objects and their related information relevant to the visualization process. These databases usually are of different classes data. These may be classified by their data types into five principal types:

Vector Data: The representation of spatial features by the definitions of their geometric forms as points, lines, polygons, surfaces and solids.

Raster Data: A representation of spatial features as digitized images. This is the typical format for still images, either arial photos or remote sensed images, and video - a contineous recorded digitized image of some geophysical phenomenon.

Structured Text: Organized text or attributes of spatial data that describe characteristics of spatial features. These are defined as records with fixed or variable length fields of text, symbols and icons.

Free Text: Free formatted text in the form of documents. e.g., treaties, reports, land titles or deeds, etc.

Knowledge Base: Recorded facts and derivation rules for logically deducing further facts and making logical inferences.

Albeit not every GIS application requires all five classes of data-types, a GIS application normally requires data from two or more of the above classifications. One problem normally encountered in GIS today is that the available data is generally not in the format required by the user. Some processing and restructuring are required before it can be incorporated into the user's application environment. About seventy percent of the time to develop a GIS application, is devoted to structuring, organizing and setting up the database. The rest of the time is divided between developing the user interface (about 20%) and the analysis and display (about 10%). The latter activities are aided by the software tools and libraries provided by the GIS software.

Naturally, the time and effort spent on reformatting the data can be reduced if the data is already available in a format that is readily utilized. The phrase *making data GIS ready* has been coined for this effort. This has driven organizations to take seriously, the development of standard exchange formats. Some of the well known initiatives include the development of DIGEST [9], SAIF [20], and SDTS [22]. One still questions if the efforts being put into the development of such exchange formats are really justified?

Consider an alternative view of the problem of making data GIS ready. We recognize that nearly all GIS software, e.g., ARC/INFO, SPANS, CARIS, MAPINFO, etc., retain spatial data in some proprietary data structure and then manage the related attribute information with a database management system (often, a relational DBMS such as INFO, Oracle, Ingres, etc.). The main reason for maintaining the spatial data in a proprietary format, it is argued, is for performance efficiency. Given that the spatial data can be maintained as a set of relational tables, it is clear then that the total information within the GIS can still be maintained by a database management system. If for performance reasons, a vendor requires, proprietary format to store the data, this can easily be retrieved from the database and restructured into the vendor specific GIS software environment. Such an internal representation is normally irrelevant to the GIS users.

Our main thesis in this paper then is that the problem of data interchange between the GISs is really that of data exchange, or interoperability between database management systems. This alternative view of data interchange problem in GIS is the motivating factor behind the development of a Federated Spatial Information Management System code named Delta-X.

The current scenario in a GIS application development is that the required data sets are available from agencies and institutions that, for operational reasons and the nature of their organizational structures, are geographically dispersed. Development of value added spatial data sets, by different organizations, sometimes are referenced to different base maps or to base maps that are out of date. The data purchased by users from data producing agencies follow a zero-one function. For example, in making a request for some selected features of a specific map sheet, the user must purchase all the information compiled for that map sheet and extract the feature relevant to his applications. The tendency then is for users to avoid the cost of repeated purchases of the same digital information. A digitized map sheet is purchased once and reused over a long period even when the data is out of date. The different rates in which different organizations compile, and revise their data sometimes force organization that are very dynamic to collect the spatial data they require themselves instead of waiting for a copy from the authoritative source that may be very slow paced in its operations. The result is the proliferation of inconsistent replicas of the same digital information at different precision and accuracy.

Granting that GIS software is highly dependent on the underlying databases, it appears then that some of these problems can be easily resolved by providing an integrated but selective access using a distributed database solution. Unfortunately, the scope of the problem is still beyond that addressed by a distributed database management system [3, 4, 13]. Rather, a multi-database system [3, 13, 17], has the potential of resolving a greater proportion of the problems. Some problems not easily addressed include:

Issues of Legacy Data: It should be recognized that the compilation of most GIS data predates the development of database technology. As a result, large volumes of datasets are currently not maintained with any DBMS. They are maintained simply as large sequential files on reels of tapes.

Underlying Data Models of DBMS: Even if all the data can be managed using database technology the databases have independent custodians and as such, have underlying database management systems (DBMS) that are based on different data models: relational, network, and object-oriented.

Cost of Adopting a DBMS: The volume of these data-sets requires a major financial commitment and undertaking to restructure them using DBMSs. Only a small percentage of the data can often be feasibly made available on-line. In some cases the rate of data capture is so high that the responsible agencies have no chance of ever restructuring the already existing data.

Data in the Past: Most database management systems operate with very little support for the time dimension. Unlike the corporate databases, e.g., personnel records, GIS users sometimes still require access to both past and current data to model and predict situations in the future. Efficient access to temporal databases requires support from the underlying DBMS. Most commercial DBMSs do not adequately support temporal queries.

The issues raised above suggest then that a solution to most of these problems requires the development of a spatial multi-database that provides efficient access to both legacy and temporal databases [21]. Strategy for interoperability in a network of database management systems, that use the same underlying relational data model, has only recently been addressed [8, 10, 5]. The question of heterogeneous database interoperability is still an active area of research [1, 13, 16]. In the mean time, there is a large number of operational GISs being used in different circumstances and in different applications. How can the contents of these multi-databases, be accessible from different Geographic Information Systems?

The Federated Spatial Information Management System (FSIM) attempts to address this question in a consistent and coherent manner. Basically, it is a multi-database system with a common integrated global conceptual schema definition. The global database is intended to achieve interoperability between DBMS that have been specialized for three distinct data-types: spatial data (i.e., vector and raster), structured text and free-text. Our design provides an integrated access to data stored in relational databases, object oriented databases, simple file systems and information retrieval and document management systems.

The current implementation does this through *Remote Procedure Calls* by formulating queries in the format of the target systems. Where the target systems are relational DBMS, SQL is used as the language for remote data access. Simple file systems are mapped into relational tables. By this means FSIM, establishes interoperability between relational DBMSs and GISs. We are currently conducting research to develop a single global database model that integrates both the spatial, structured attribute and free-text models into a common object-oriented database framework.

This paper describes some aspects of the FSIM implementation that is code named Delta-X. In particular, we discuss some details of the implementation and the rationale for our design choices. In the next section, we present some of the design problems and other related research works. In section 3, we present the general configuration framework of Delta-X. In section 4 we present some fundamental requirements and the approach to implementing the federated spatial multi-database system. The main functionalities of Delta-X are discussed in section 5. Some problems beyond those answerable by technology but essential to the operations of Delta-X are raised in section 6. In section 7 we conclude and discuss direction for future work.

2 Problem Specification

Consider the situation in a local area network in which a number of workstations (nodes), each runs a different vendor GIS software. Due to the different internal format for representing information, each workstation must maintain a separate copy of the data it works on. Ideally, the workstation should operate as a client of some database server that maintains a complete repository of all the data relevant it. Rather than maintaining multiple copies of all the data sets (e.g., map sheets), the server should maintain all the databases and a user at a workstation *checks-out* and *checks-in* the required data-set that he or she works on.

The preceding discussion reflect only a subset of a much wider problem in GIS. Users of GIS technology can be categorized into four groups: technology developers, data producers, application developers and management and decision makers. We concern ourselves with the data producers and application developers (or data consumers) since these form the community of GIS users that manipulate the database in a substantial manner. The management and decision makers normally interact with the data in *read-only* mode. The forgoing illustration of heterogeneous GISs in local area network (LAN) can easily be perceived in much wider scope of a global information network, see figure 1.

Data producers and consumers are all geographically dispersed. If the data produced is to be made available to the user in a timely manner, then the data producers and consumers must be interconnected in a wide-area-network. Given a wide area interconnected network of data repositories, the average GIS application developer at a workstation, requires two main services: a mechanism for easily identifying the data resources required for his or her application and a mechanism for accessing and transferring the data to his GIS environment.

Delta-X address the latter problem from two standpoints: The first is a service in a heterogeneous GIS local area network having a data server. The second is a service in a wide area network of data servers. A typical WAN and LAN configuration is shown in figure 1. To highlight some of the Delta-X features, we assume that each GIS workstation is configured as a database client of some database server. Basically, the interaction between the GIS workstation and the server, with respect data accesses, is based on a *client-server* architecture [14, 18]. This model of operation reflects the situation in most organizations. The problem is complicated by the fact that the data types to be retained by the servers are varied.

A typical interaction of an application developer with a GIS, consists of:

- Checking data out of the server and restructuring it into its local data storage.

- Developing a user interface tailored to the application.
- Displaying the results of his or her analysis in different visualization modes: graphics, photo-imagery, text, animated modeling, video, etc.
- Checking the data back onto the server, either as a new version or as an updated and revised form of an existing one.

This mode of operation raises a number of questions.

- Given the different data types, what database management system can retain, in a consistent manner, all the data of a server?
- How should updates be reflected in the stored data? In particular, is there a storage efficient method for managing revised copies from which past versions can be reconstruct.
- Given the information on the server, can a user request data over an area for only a selected set of features? What concurrency control mechanism would be most efficient for multiple data access under such mode of operation?
- In the case of data held on special devices, e.g., CD-ROM, can information be concurrently accessed by multiple user?
- What is the optimum data capacity for a cost-effective on-line data distribution given users' access profiles?

These and similar related questions are what the Delta-X project is concerned with. Such a model of operation is not exclusive to GIS. Similar problems have been addressed in purely corporate database environments. One approach to managing heterogeneous databases is the use of multi-databases management systems [3, 11, 13, 17]. Except for, sound (i.e., speech, voice and music), GIS data requirements are similar to those of multimedia computing. Similar problems are encountered in a network of multi-media databases [6].

On a wide area network of GIS data servers that service thousands of clients, each data server must be configured for very high rate and long lived transaction since typical operations involve the retrieval and restructuring of data. This may take some considerable amount of time to complete. As in the case of multi-media database, the potential for high volume data transfers requires very high network bandwidths, good data compression and decompression techniques with no information loss, large storage capacity as well as high performance I/O bandwidth between main memory and secondary storage.

Storage and delivery of data to client workstations constitute only one aspect a series of services in Delta-X. Other services include administrative support, secure and reliable data capture, delivery of data to a large customer base, visualization of results of remote computations. We describe the basic ideas of Delta-X, its functionalities and its current implementation status in the subsequent sections. A number of projects and researchs are being pursued elsewhere to address similar problems. These include the Geo-DASDBS [16], the *SEQUIOA* project [19] and the *Papyrus* project [12].

3 A Federated Spatial Information System

3.1 Basic Configuration

To establish the mode of operation in network of GIS workstations as discussed in the preceding sections, a federated spatial information management system development (Delta-X) was initiated in the GIS Division, Natural Resources, Canada. Delta-X is best characterized as a federated multi-database system [3, 17] that is constrained for:

- remote *read-only* data access,
- remote computations,
- remote display and
- limited database management system functions.

The global model is materialized, using a set of consistent transformation rules, onto local databases. The local database management systems (LDBMS) that can be relational, object-oriented or a simple file system. Similarly, the schema definitions of a local database may mapped to the global database conceptual model. Our present implementaton suports databases that must be relational or simple file systems. However, we will assume a global data model that is based on *extended-entity relationship* data model. Unlike a general multi-database system where the problem of schema integration is a much complex issue to handle, this problem does not arise in our constrained multi-database configuration. The global data model is such that it can be mapped onto either an underlying relational DBMS or an object-oriented DBMS.

A typical configuration consists a network of independent autonomous databases management systems that form the nodes of the Delta-X servers. These form the local database management systems (LDBMS) of the Delta-X federated database management system. The clients on which the GIS softwares run on are grouped into client nodes of a LAN. A LAN will data server that runs the Delta-X server component software (*DxServer*). A *DxServer* runs on a capable Unix system that communicates with other *DxServer* in a wide-area-network.

The Delta-X services are split into two components: the server services (*DxServer*) and the client services (DxClient). The Delta-X global data model integrates spatial (vector and raster), non-spatial data, free text and documents. We anticipate incorporating a knowledge base but this is not addressed in the current implementation. The defined global schemes of the respective data classes are virtual and materialize only as the stored data of the corresponding local schemes. The global schemes are therefore referred to as *virtual.* However, the global schema definitions are supported in a Delta-X directory information service (*DxDIS*), that is fully replicated at every *DxServer* node.

The concept of a federation is established at the database level. A database administrator (DBA), manages one of the *DxServer* nodes. The DBA is responsible for registering databases of an LDBMS into the federation. Registering a database implies exporting the schema definition of a local database to the global database of the DxDIS. A Delta-X master node (DxMaster), maintains a master copy of DxDIS. This node is maintained currently, by the server of the GIS Division, Natural Resources, Canada. The DxMaster node administers the functions of admitting and retiring Delta-X servers (and not the databases). Once a server is admitted, into the federation, the DBA of the *DxServer* freely registers and de-registers databases maintained by the local DBMS.

3.2 A DxServer Node

Due to the different data types required in a GIS, a *DxServer* may be configured to maintain and deliver more than one data type. For example, a server may be configured for only vector data that is topological structured. Another may be configured to manage either only remote sensed data, a national toponomy database, geological database or hydrographic data. A *DxServer* node runs cooperatively with a client node during data translation to the respective GIS format.

Transaction services of the *DxServer* will be generally long-lived. If the anticipated number of access is sufficiently large, the *DxServer* must be computational powerful to handle the workload. A *DxServer* node may be a super-computer, a highend graphics engine or a cluster of workstation that perform parallel and distributed computation using virtual parallel machine (PVM) tools [7]. In some case, due to the specialized nature of data processing activity, availability of special skilled personnel, the operations of the server may be dedicated to special operations. An inexpensive manner to support a *DxServer* is to use a cluster of workstations that pool their processing power together using the parallel virtual machine paradigm.

3.3 Remote Data Access and Queries

A DxClient runs on a Unix workstation that is configured also as an X11R5 server with PEX extensions and a Motif GUI interface. The DxClient is scheduled to be ported to a PCWindow's environment. Normally, a DxClient workstation is also configured to a GIS software and as such, will have the required graphics support needed. A DxClient node has connectivity to some *DxServer* node. Clients perform remote data access either directly using a graphical user interface of the DxClient component from the client node or by explicit query specification using a common global Delta-X Data Access Language (DxDAL). The GUI interface mode of data access is the simplest. This will normally be the mode for public usage [14].

Using this mode of interaction, a client interrogates its default server for the collection of databases available. This collection consists of the databases globally available and accessible to the user and those residing in the local database of the client's server. The client may request that some data be retrieved and made GIS ready in the client's environment. This is done by simply specifying the source data set name, the required features and the destination of the data.

For example, from a set of vector formatted map sheets, a user may request only specific features of a particular map sheet. The result of such a request is the digital vector data of the same map sheet that is topologically structured and restricted only to the selected features. The unit of data access is a collection of selected features from a specific data set.

The *DxServer* to which a client is currently connected may not necessarily execute the service requested by the client. The transaction may be forwarded to the appropriate *DxServer* with the result forwarded directly to the client.

Advanced users may utilize a lower level query facility in which the data access is done by specifying a query in the common data access language (DxDAL). The common data access language in the current implementation is a modified SQL for multi-database access (MSQL). A user in a DxClient environment, has access to the DxDIS database and the local database directories or catalogues. A query is expressed in MSQL with reference to a global table as *"database_name@server_name.tablename"*. Access using the GUI is still executed indirectly by the underlying MSQL.

In its simplest form, a query in MSQL is issued from a client to the connected *DxServer*. This is then analyzed and decomposed into subqueries that are destined to the respective LDBMS holding the data repositories referenced in the subqueries. A *DxServer* agent of the receiving LDBMS, translates its subquery into the specific local database query. This is then passed on to the local DBMS for execution. The results of the respective subquery execution are then passed to the coordinator (the server originating the query on the client's behalf) who then assembles the result for the requesting client. All remote queries are executed as *read-only* transacions. Updates to database are carried out only at the local database management system. This is not a limitation but the desired mode of operation since the autonomy and independence of the *DxServers* must be respected.

Since remote client requests are read-only, distributed two-phase commit is not an issue in the model of distributed access. Remote queries that arrive at a local node is transformed to a local tranasction that is excuted concurrently with other local transactions. Hence a simple timeout mechanism is sufficient to coordinate multidatabse transactions and maintain database consistency.

4 Common Spatial Data Model

4.1 The Delta-X Approach

Essential to the mode of operation is a common data representational scheme. This is defined in Delta-X by the common global conceptual data model. Consider the problem of providing topologically structured vector data to different GIS clients. Delta-X maintains a common global database scheme which may or not be adopted by any inndividual GIS as its representational data model. This is the conceptual data model into which other GIS internal representations can be mapped without information loss. The Delta-X spatial data model, for vector data in this case, maintains up to degree *degree four topology*. This global scheme is considererd to be *virtual* since no explicit global data representation exists other than the database of the directory information service.

This global conceptual scheme forms the intermediate transition schemes for data exchange between different GISs. The global conceptual data model has a mapping onto either a relational database, as a collection of relational tables, or into an object-base as a *colony* of *categories.* A *DxServer* maintains the vector data in a relational database if the underlying database management system is relational, e.g., Ingres, Oracle, Informix, etc. Similarly, the *DxServer* retains the data in a set of equivalent category classes in an *object oriented database,* if the underlying DBMS is object-oriented, e.g., ODE, ObjectStore, Objectivity. The significant idea is that *DxServer* transparently delivers data to and receives data from clients' GIS environment.

The ilustration with the vector data shows the general approach for handling data in Delta-X. This approach, where by a common defined global data model is materialized for actual representation in local databases management system, is extended to handle the various classes of datatype required in GIS. Namely, raster data such as the arial photo and remote sensed images of point, line and area features, structured text that are related to spatial objects, free text, digital elevation models (DEM) of areas, etc., are all represented in the common global schema definition. Since some of these data types have internationally defined standards of representation, e.g., GIF, JPEG, MPEG, SGML, HTML, etc., these are maintained as files in their respective standard formats, and related to the defined spatial features through spatial indexes. A number of such non-vector data types are related to the spatial features using index techniques.

4.2 Spatial Indexing

Even over a small area of coverage, maintianing all the relevant information for all data types, quickly grows into a siginificant large database. The density of information is generally non-uniform. For example, sparsely populated areas have less accumulated data than densely populated areas. Delta-X organizes space into hierarchical tesselated regions and maintains two distinct levels of index schemes for identifying a feature in defined space. The first level of index identifies a regular polygonal cell of the region of coverage that has been hierarchical tessellated into a near uniforn grid. Tbe second level of index associates features within each cell with other informtion types. The first level uses a quadtree-like method of spatial index [15]. The second level draws from a number of one-dimensional and multi-dimenesional index techniques depending on the data-type to be indexed.

The partitioning of the space into cells (referred to as tiles in the literature), by the first level of index enables the otherwise large volume of data to be accessed and controlled in manageble units. By this means, all data of a specific region or cell, and controlled by a particular organization or agency, can be grouped together. Geographic coordinate system (longitude and latitude) forms the common and basic coordinate reference system by which regions and features are related across multiple databases. The local database may derive the actual projection and coordinate reference system used in referencing spatial objects within that particular environment.

5 System Functionalities

In indentifing the functions of the federated spatial information management systems it is important to note that *Delta-X* is not a Geographic Information System. It is only concerned with the facilities for managing a network of database that are to be integrated in a GIS application.

The functions desired in such a system is better percieved from the stand point of users, system administration, and policy makers. The desired functions from a user's standpoint depend on whether the users are data providers or data consumers. Data providers generally require to administer a *DxServer* which they have control over. They can then make their databases available to others, namely researchers and application developers. Since some cost is associated with the compilation of these databases, data providers would normally need some return of their investiment. Hence some accounting and billing functions are expected. The data consumers require functions that assist them to locate readily the databases relevant to their applications and to retrieve the data into their GIS environment in some desired format.

5.1 Basic Functions

From a local database management system point of view a *DxServer* is a layered software on top of the LDBMS that acts as the agent of Delta-X. From this standpoint, *DxServer* functionalities include:

- registering and de-registering databases;
- translating user queries to and from the underlying LDBMS forms;
- transfering data between one server and another; i.e., *DxSever-to-Dxserver* transfers;

- cooperating with specific vendor GIS software to translate retrieved data into the GIS specific format;
- monitoring and informing clients of the status and progress of requested service;
- logging the relevant information about clients, for accounting and billing purposes when the data being accessed has a cost associated.
- authenticating authorized users and exercising security measures.
- forwarding requests or migrating transactions, to other *DxServers* where the server cannot process a request or complete the transaction;

The above discussions give some of the services carried out by a *DxServer*. From the *DxClient* standpoint, a user invokes operations either using the graphical user interface or the common query language (MSQL). The functions achieved through the *DxClient* include:

- request for the transfer of data-set between a client's local *DxServer* and a *DxClient* workstation, i.e., *DxServer-to-DxClient* transfers and *DxClient-to-DxServer* transfer;
- selective data retrieval by features;
- incremental visualization of spatial features in different modes: vector maps, photo-images, remote sensed data digital elevation models, approximation models, etc.;
- seamless horizontal and vertical spatial navigation;
- display of attributes, free text and documents associated with spatial features;

There are a number of functions that require cordinated support between the local database management systems and the multi-database management systems. These are normally not initiated from users at remote clients but as database administrator of the respective local *DxServers*. We refer to these as special administrative functions. They are functions such as:

- monitoring and maintaining data access profile for subsequent performance tuning;
- billing users for data that is accessed at some cost;
- data revision and control for efficient access to information in the past and processing of temporal queries;
- granting user authorization to access information and protecting against unwarranted data access;
- recovering from system failure;
- validating the consistency of data, particularly with respect to different referencing mechanisms at different servers;

5.2 Global Policy Administration

Ideally, one would prefer an operational system that is self administering with respect to enabling and disabling the local databases from anyone participating *DxServer*. One central control site (*DxMaster*) overseas the admission and departure of sever nodes into the federation. Unfortunately, there are still rules that must be discussed and agreed upon by participating members in the federation. Therefore, some policies must be established servers administrators must make every effort to adhere to these policies for efficient and non-disruptive service to clients.

While the autonomy and independence of the data sources for GIS application development must be respected, some common agreement must be reached as to whom has the authoritative control of a particualr data set. For example, it is possible for one agency to purchase data from some source, restructure the data into a cleaned data with some added values and enhancements and make it accessible to the public as a better alternative to the original data. Such an action could be detrimental to the prestige of the original provider of the data. Matters of this nature and other legal issues need to be addressed as policies of participating member in the federation.

6 System Status

Work on Delta-X is still in progress towards implementing a full functional system. Our current prototype, operates with three *DxServer* systems: two relational database management systems that execute SQL queries, Ingres and Oracle, and one simple file system. Several GISs run DxClients that interact with these servers. The frontend, GUI, of Delta-X is completed. The first level display window on invoking Delta-X is showm in Figure 3. Figure 4 shows the display of a vector map from a remote server. The implemented systems rely purely on remote procedure calls for its remote operations.

The Delta-X project began before serious initiatives with respect to standards for interoperability between commercial relational databases were fully specified. Work on standards for interoperability in object-oriented database is still ongoing and for standards for interoperability between multi-databases is yet to be drafted. We have since identified close similarity between the multi-database requirements in Delta-X the functional services provide by the Xopen reference model for distributed transaction systems. We are exploring the use of a commercial distributed transaction that meets existing international standard and satisfies our current operational functions. This, we believe, can be easily adopted to reduce our development time for the multi-database transaction modules of our system.

7 Conclusion

The Delta-X operational paradigms describe a methodology for the integration of multi-database management concepts and data interchange in a network of heterogenous databases. The system as designed is a sufficiently open system that freely admits new participants, either as new databases management systems or new GIS technology, with minimum overhead. The only essential requirement is that new database or GIS vendors must provide the mapping functions that translate information between their propriatory data format and our global conceptual data model.

We have described the essential functionalities in Delta-X and we have shown the software modules required to be integrated to achieve a successfull implementation. GIS database development is an exercise being carried out as national projects in several countries. We believe the approach taken by the Delta-X design will form the reference model by which similar projects will emerge to establish eventually, a global network of GIS databases.

References

[1] R. et al Ahmed. The pegasus heterogeneous multidatabase system. *Computer*, 24(12), Dec. 1991.

[2] S. Aronoff. *Geographic Information Systems: A Management Perspective.* WDL Publ., Ottawa, Canada, 1989.

[3] D. Bell and J. Grimson. *Distributed Database Systems.* Addison-Wesley, Co., 1992.

[4] S. Ceri and G. Pelagatti. *Distributed Databases: Principles and Systems.* McGraw-Hill, Inc., 1984.

[5] X/Open Corp. X/open database language specification. Specs. Standard, 1991.

[6] E. Duval, H. Olivie, P. O'Hanlon, D. G. Jameson, N. Ismail, S. Wilbur, and R. Beckwith. Managing networked multimedia data. *Computer Graphics*, 28(1):15 – 19, Feb. 1994.

[7] A. Geist, A Beguelin, J. Dongarra, W. Jiang, R. Manchek, and V. Sunderman. *PVM 3 users's guide and reference manual.* Oak Ridge National Laboratory, Oak Ridge, Tennessee, 37831, May 1993.

[8] J. Gray and A. Reuter. *Transaction processing: Concepts and Techniques.* Morgan Kaufmann, San Mateo, California, 1993.

[9] Digital Geographic Information Working Group. *DIGEST: Digital Geogaphic Information Exchange Standards.* Draft Report, Enviromental System Research Institute, 380 New York Street, Calfornia, Sept. 1990.

[10] SQL Access Group. Sql access formats and protocols (fap) specification. Specs. Standard, 1991.

[11] D. Heimbigner and D. McLeod. A federated architecture for information management. *ACM Trans. on Office Information Systems*, 3(3):253 - 278, Jul. 1985.

[12] C. P. Kolovson, M.-A. Neimet, and S. Potaimianos. Interoperability os spatial and attribute data manager: a case study. In D. Abel and B. C. Ooi, editors, *Advances in Spatial Databases*, pages 239 - 263, Springer Verlag, Jun. 1993.

[13] W. Litwin, L. Mark, and N. Roussopoulos. Interoperability of multiple autonomous databases. *ACM Comput. Surveys*, 22(3):267 - 293, Sept. 1990.

[14] J. Medved and G. Petras. Client-server architecture of delta-x. In *In this proceedings*, Jun. 1994.

[15] H. Samet. *The Design and Analysis of Spatial Data Structures*. Addison-Wesley, Reading, Mass., 1990.

[16] H. J. Schek and A. Wolf. From extensible databases to interoperability between multiple databases and gis applications. In D. Abel and B. C. Ooi, editors, *Advances in Spatial Databases*, pages 207 - 238, Springer Verlag, Jun. 1993.

[17] A. P. Sheth and J. A Larson. Federated database systems for managing distributed heterogeneous, and autonomous databases. *ACM Comput. Surveys*, 22(3):183 - 236, Sept. 1990.

[18] A. Sinha. Client-server computing. *Comm. ACM Journal*, 35(7):77 - 98, Jul. 1992.

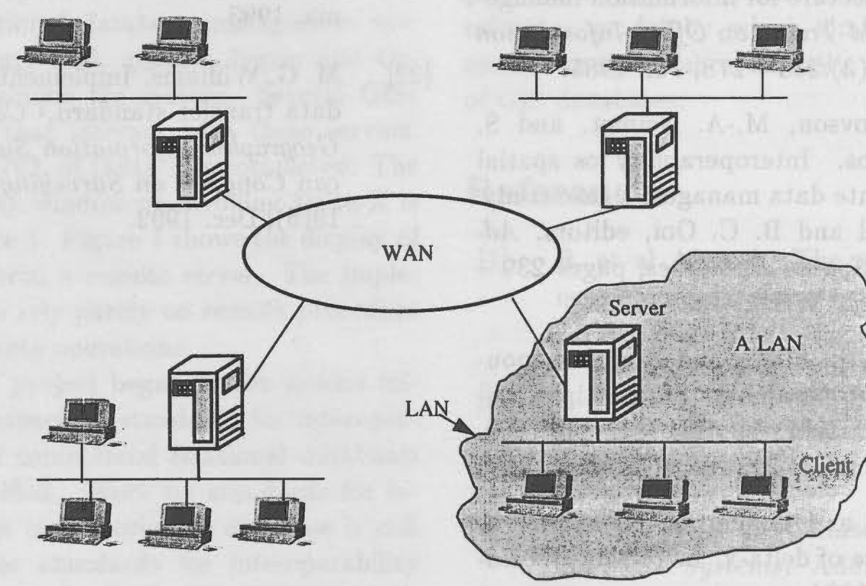
[19] M. Stonebraker, J. Frew, and J. Dozier. The sequoia 2000 project. In D. Abel and B. C. Ooi, editors, *Advances in Spatial Databases*, pages 397 - 412, Springer Verlag, Jun. 1993.

[20] Surveys and Resource Mapping Branch. *SAIF: Spatial archive and interchange format: Formal Definition.* Province of British Columbia, Aug. 1992.

[21] A. U. Tansel, J. Clifford, S. Gadia, S. Jojadia, A. Segiv, and R. Snodgrass. *Temporal databases: Theory, design and implementation.* The Benjamin/Cummings Publishing, Co., Inc., Redwood City, Caifornia, 1993.

[22] M. G. Williams. Implementing the spatial data transfer standard. *Cartography and Geographic Information Systems: American Congress on Surveying and Mapping*, 19(5), Dec. 1992.

Figure 1: A Wide Area Network of GIS LAN and Servers



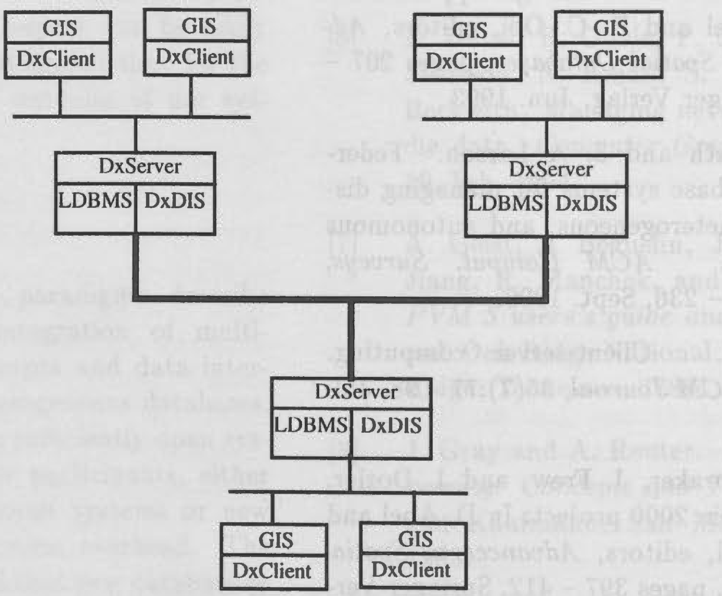


Figure 2: A Logical View of Delta-X Client-Server connection

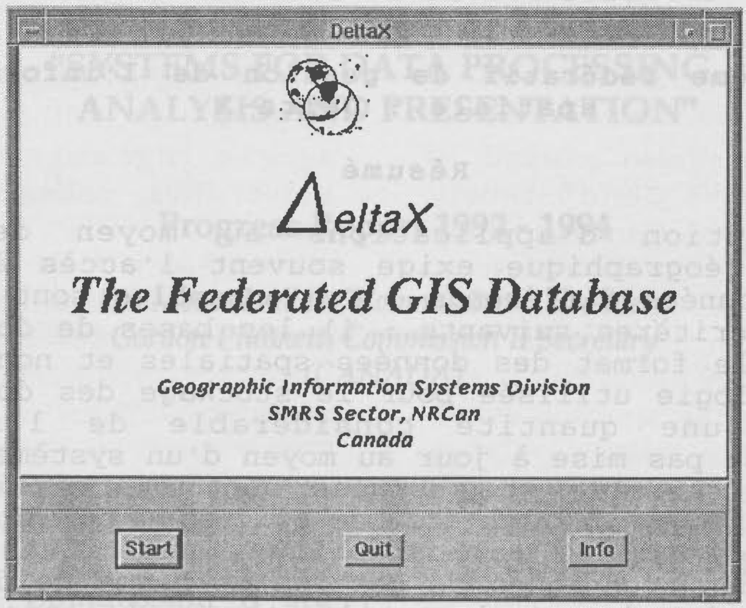


Figure 3: A frontend of Delta-X

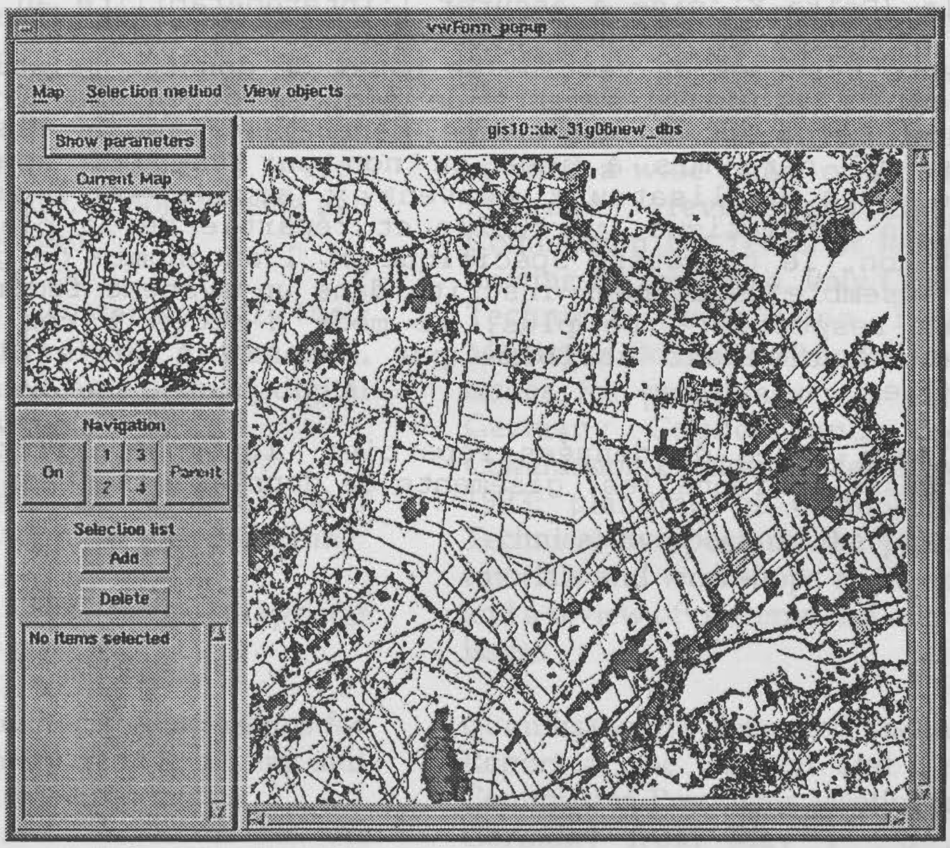


Figure 4: The Main Client Display Window of Delta-X Server

Un système fédératif de gestion de l'information spatiale : Delta-X

Résumé

La conception d'applications au moyen de systèmes d'information géographique exige souvent l'accès à plusieurs ensembles de données différents. Ces ensembles sont hétérogènes au regard des critères suivants : i) les bases de données sous-jacentes, ii) le format des données spatiales et non spatiales, iii) la technologie utilisée pour le stockage des données, etc. Actuellement, une quantité considérable de l'information nécessaire n'est pas mise à jour au moyen d'un système de base de données. Les ensembles de données sont stockés et transmis simplement sous forme de fichiers, ce qui peut être dû au fait que la collecte de données cartographiques numériques a précédé l'élaboration des techniques de gestion de bases de données. Il en découle que la plupart des utilisateurs obtiennent de façon indirecte l'information qui leur est accessible. Par conséquent, les systèmes d'information spatiale, et les SIG en particulier, ne permettent pas un accès intégré et unifié à la multitude de données de types divers qui sont nécessaires à la conception d'applications. Le système fédératif de gestion de l'information spatiale (Delta-X) vise à assurer l'interopérabilité au sein d'un réseau de bases de données hétérogènes. Il fait le lien entre différents systèmes de gestion de bases de données relationnelles, des fichiers et des gisements de données orientées objets. Au moyen du Delta-X, un organisme peut inscrire un ensemble de données qui sera mis à jour ou non par un SGBD sous-jacent. Ensuite, des utilisateurs peuvent y accéder à l'aide d'une définition schématique, virtuelle et répartie de la planète. De cette façon, le caractère spécifique et l'autonomie de gestion de chaque ensemble de données inscrits dans un système fédératif sont toujours respectés. De plus, le mode d'accès à ces ensembles demeure relativement uniforme par rapport à un environnement particulier.

Mots-clés : bases de données réparties, interopérabilité, modèle de données relationnelles, gisements de données orientées objets.

ISPRS TECHNICAL COMMISSION II

"SYSTEMS FOR DATA PROCESSING, ANALYSIS AND PRESENTATION"

Progress Report 1992 - 1994

Dr. Mosaad Allam, Commission II President
Gordon Plunkett, Commission II Secretary
(CANADA)

TERMS OF REFERENCE

The interests of Commission II are dedicated to the following activities:

- design and development of integrated systems for measurement, processing, analysis, representation, and storage of photogrammetric, remote sensing and GIS data
- study and evaluation of system integration aspects for photogrammetry, remote sensing and GIS data processing
- analysis of systems and their components for automated, semiautomated and manual digital processing systems
- development of systems and technologies for radar data processing
- study of real-time mapping technologies
- standardization of digital systems for photogrammetry, remote sensing and GIS

ACCOMPLISHMENTS OF THE COMMISSION DURING 1993

Commission II held a chairpersons meeting during the 5th Annual Canadian Conference on GIS from March 27-29, 1993 in Ottawa, Canada. The Conference provided the opportunity for a meeting between the chairpersons of three Working Groups, and co-chairpersons from several other W.G.s and Commission II's Executive. The meeting covered a wide range of topics which included: review and finalize of Commission II's Terms of Reference; a briefing on the Stresa, Italy Council and Technical Commission Presidents Meeting of October 1992 and discussion of issues for presentation at the Bonn, Germany Council and Technical Commission Presidents Meeting of May, 1993; planning of Commission II's Technical Symposium; the planning of tutorials and workshops at Conferences; and a review of concerns and outstanding issues.

Commission II presented its mid-year Report at the Council and Technical Commission Presidents Meeting in Bonn, Germany from May 4 - 5 , 1993. Commission President Dr. Mosaad Allam represented the Commission and forwarded all Commission II concerns and outstanding issues. Some modifications were made to the Terms of Reference of several Commission II Working Groups.

Commission II has also worked to establish effective and timely communication with Working Group chairpersons and co-chairpersons. Throughout 1993 the Commission has diligently written, updated and distributed a host of information and material related to the activities of the ISPRS. In response, Working Group chairs and co-chairs have replied in an equally effective fashion. In particular, the Commission strives to use electronic means whenever possible.

During the year of 1993, Commission II has worked extremely hard in preparation of its Technical Symposium from June 6 - 10, 1994 in Ottawa, Canada. Associated with these preparations have been a number of activities, which include: forming a Technical Program Committee to oversee all aspects of the Technical Program; forming an Accompanying Persons Committee to prepare an active and exciting social calendar; distributing two Call for Papers (see below for details); assembling the abstracts received in response to the call for papers; preparing the Preliminary and Final Programs; solidifying logistical activities related to the Symposium i.e. reserving the exhibit and session areas, hotel arrangements etc.; and collating the Symposium Proceedings.

Distribution of the first Call for Papers in March of 1993. Distribution of a second glossy Call for Papers was distributed in September of 1993. In each mailing of the Call for Papers were sent to over 17,000 addresses across the world.

A large colour poster (measuring 2.5 x 3 feet) advertising Commission II's June Symposium was printed and distributed to over 1000 agencies related to the areas of ISPRS, including: large national agencies, large and medium sized firms, universities, colleges and research institutes.

Commission II's Executive held a meeting with Mr. Jacques Kiebert of Elsevier Publishers (producer of the ISPRS Journal of Photogrammetry and Remote Sensing). The meeting was most cordial and explored possible areas of cooperation, including preparing special theme journal issues, expanding the journals distribution and exchanging advertising of activities.

Contacts with Dr. David Tait, Editor of the ISPRS Journal for Photogrammtry and Remote Sensing have also been initiated. A proposal for theme issues of the Journal has been presented and plans for future cooperation between Commission II and the ISPRS Journal are being explored.

STATE OF SCIENCE AND TECHNOLOGY COMMISSION TOPICS

As changes in the applications of data continue to expand, so does the expansion of systems technology for the handling of data. Together these two issues dominate the areas of interest for Commission II. The rapid evolution from manual time-intensive processing and analog analysis systems to faster more precision based computerized systems has led to a greater appreciation of the applications of the data. And, greater applications of data have not only expanded the community of users and interested parties, but have also led to greater linkages with other disciplines and scientific based studies.

These have and will continue to only strengthen the position of photogrammetric and remotely sensed data by forging greater alliances and markets for future data and expanded applications.

It has also become apparent that each improvement in using data in a system, invariably leads to an improvement in efficiency, especially those related to human and financial resources. While it can be argued that efficiency is not always achieved and not always in a cost efficient manner, it is a paramount consideration for the future, in terms of both data volumes and availability. Increases in operating and production costs are initial, and will lead to greater long run reductions in cost and improvements in efficiency. However, the balance of this reasoning is threatened by the absence of marked international standards. Standards for data, both source formats and outputs, and systems. This is a concern which needs attention and should be approached within a structured framework, as the ISPRS.

COMMISSION II NEWS

June 4 - 5, 1994 - a wide range of tutorials and workshops are scheduled in advance of Commission II's Symposium opening. These workshops will be half or full day in length and cover many topics. At present, 12 different workshops are planned, some involving hands-on training others in a presentation style. Among various others, a sample of titles includes:

" Putting Remote Sensing Imagery through its Paces: Data input, Analysis, Map Output, and use in a GIS"

"GPS for Photogrammetry and Aerial Mapping"

"Automatic Vehicle Location/Navigation/ Guidance Systems (AVLNG)"

"Terrain Mapping Using Synthetic Aperture Radar"

"Mobile Data Collection for GIS"

June 6 - 10, 1994, Commission II mid-year Technical Symposium in Ottawa, Canada. A full technical program of paper, poster and plenary sessions will highlight the research activities of Commission II.

A large and exciting Exhibitors area of over 60 booths was organized for Commission II's Symposium. The exhibit area is shared with the Sixth Annual Canadian Conference on GIS, and will provide the opportunity to view and access the most advanced technology and private sector research developments. A forum to provide presentations on technology is also scheduled.

A strong and dynamic accompanying persons program for Commission II's Symposium has also been planned. Included among these activities are tours of the city, visits to National museums, social cocktail parties and evenings of fine dining. A spectacular awards banquet has also been planned.

The June, 1994 Commission II Symposium also presents an opportunity for an ISPRS Commission II Executive meeting, which is scheduled for Wednesday, June 8 from 3:00 - 5:00 with ISPRS Council members attnding.

WORKING GROUP ACTIVITIES DURING REPORT PERIOD

WG II/1 - "Real Time Mapping Technologies"

Chairman:	Dr. Kurt Novak (USA)
Co-Chairman:	Michael Hahn (Germany)
Secretary:	Holger Schade (Germany)
WG Members:	80

Terms of Reference

- design and development of real-time mapping systems
- survey recent and future developments in real-time mapping systems
- design of integrated digital systems for real-time utility mapping and GPS van technologies
- investigate sensor information processing and analysis issues in autonomous vehicle navigation systems
- assess the role of stereo-vision and kinematic GPS technologies in integrated real-time mapping systems
- establishment of contacts with working groups of Commission I, II, III, V interested in integrated sensor orientation, GPS-INS integration, object recognition and other relevant issues

Accomplishments of Working Group II/1

During the past year the activities of WG II/1 concentrated on organizing the working group, planning the activities for the four year period, informing potential members about these activities, and helping plan the Commission II Symposium in Ottawa in 1994. A chronological list of accomplishments follows:

Oct. - Dec. 1992 - edit the Special GPS Photogrammetry Issue of Photogrammetric Engineering and Remote Sensing, published Nov. 1993.

Feb. 18, 1993 - organized a session on GPS Photogrammetry at the ASPRS/ ACSM Annual Convention in New Orleans, USA

March 24, 1993 - organized a session on real-time mapping technologies at the Canadian Conference on GIS, Ottawa, Canada.

April, 1993 - distributed circular letter: Introduction of Working Group terms of reference and planned activities.

August, 1993 - distributed circular letter: Announcement of Working Group Meeting at Photogrammetric Week, Stuttgart, Germany, and address list.

Sept. 22, 1993 - working group meeting during the Photogrammetric Week, Stuttgart, Germany.

October, 1993 - distributed circular letter: Announcement of ISPRS Commission II Symposium in Ottawa and report on Working Group meeting in Stuttgart.

July 1993 - May 1994 - conduct a survey on the state-of-the-art in real-time mapping.

State of Science Technology of Working Group II/I Topics

When WG II/I began its activities one year ago there were only a few places that actively pursued real-time mapping. It was mostly done by researchers at the Ohio State University and at Stuttgart University. However, these original developments lead to related research in many other fields and at other institutions. A comprehensive investigation into this topic is under way and will be presented at the Ottawa Symposium. The most important activities are reported below.

Integration of GPS, INS, Laser Scanners and Digital Cameras for Aerial Mapping: GPS is used for automatic positioning and orientation of the sensor platform. With an operational, automatic ambiguity solution available, real-time kinematic GPS can be used for sensor platform positioning. Developments are under way to fully integrate inertial systems with GPS for on-line altitude determination, however, the high price of inertial units is still an obstacle. Digital or video images are collected together with the positional information. Aerial video is becoming a popular tool for taking inventory of forests and wildlife management. As high resolution digital cameras (> 2,000 x 2,000 pixels) are practically not available for aerial mapping, digital image collection systems are mostly restricted to time critical applications as well as mapping of linear features. This is one of the research areas that will need more attention in the future. Laserscanners are being used for surface reconstruct ion and DRM generation. They are in the prototype stage and are promising tools for real-time orthophoto production. There are also interesting developments going on to install digital cameras in the wing-tips of airplanes to take stereo-image pairs of powerlines. The position and orientation of the airplane is determined by 3 GPS antennas.

Vehicle Navigation Applications: a number of infrastructure mapping systems have been developed over the past years. They are based on the integration of GPS and INS for positioning vehicles continuously. Different imaging sensors are available for video-logging and attribute collection, as well as stereo-vision systems for spatial positioning of objects in the environment of highways. Data collected with these systems are immediately stored in a GIS. Other research deals with autonomous vehicle navigation. Image sequences are analyzed by tracking points and linear features along the road to estimate the vehicle's motion and speed. On-line triangulation techniques, such as Givens transformations, and 3D Kalman filters are applied. In this category we also include the developments in the IVHS (intelligent vehicle highway systems) field. They mostly deal with the installation of moving map displays and traveler information systems in cars and trucks, to show the position of the vehicle in real-time on a digital map.

Integration of GIS and Image Analysis: the integration of GIS data is very useful for automatic, digital image analysis. This is especially true in real-time environments, where different sensors are available that can support image analysis. Automation is a must in such an environment and can only be achieved by proper interfaces. Multi-media, spatial data-bases enhanced by different types of digital imagery provide valuable information in this endeavor.

Working Group II/I News

- Development of a report/bulletin on sensor and GIS integration is on-going. It will probably be completed for the Symposium in June 1994.
- A survey is being conducted on the state-of-the-art in real-time mapping, including the activities of different institutions and Commercial companies world-wide.
- Two technical sessions will be organized at the ISPRS Commission II Symposium in Ottawa, Canada (June 1994).

- The WG will prepare two tutorials for the ISPRS Commission II Symposium in Ottawa; topics are: "GPS for Photogrammetry and Aerial Mapping", and "Mobile Data Collection for GIS".

- A technical session will be organized at the Workshop on High Precision Navigation (Germany 1994/95).

- The WG will participate at a Joint Workshop on "Computer Vision in Photogrammetry" in Stuttgart, Germany (8-10 November, 1995); organized together with WGs I/5 and III/2 + 3.

- WG will co-sponsor a Joint Workshop with InterCommission WG II/III in USA (1994/95)

- WG will organize a Joint Workshop with WG III/I in Barcelona, Spain (1995).

- The WG will actively participate at the following conferences: Commission V Symposium, Australia (April 1994), Commission III Symposium, Germany (September 1994).

WG II/2 - "Hardware and Software Aspects of GIS"

by Chairman:	Dr. Manfred Ehlers (Germany)
Co-Chairman:	Nickolas L. Faust (USA)
Secretary:	David Steiner (Germany)
WG Members:	48

Terms of Reference

- design and operational aspects of the integration of GIS with image analysis systems

- studies of GIS characterized by workstation, mainframe and microstations in a heterogeneous environment

- studies of parallel processors, array processors, supercomputers and optical hybrid systems for improving GIS

- design and performance issues for 3-D GIS

- studies of benchmark designs for integrated GIS

- studies of man-machine interaction: displays techniques, interactive techniques and audio interaction

- hardware and software for input/output aspects of GIS

- studies of GIS standardization as applied to user interface, networking, testing and databases

Accomplishments of Working Group II/2

The working group has accomplished the following:

April 6, 1993 - Working Group Meeting at the 25th International Symposium: Remote Sensing and Global Environmental Change. The Meeting was attended by 19 participants and convened by Manfred Ehlers, who gave a brief introduction to the structure of ISPRS and its working groups. A short description of the other Commission II working groups was presented. This was followed by a description and explanation of WG II/2's terms of reference.

Workshop Planning - It was announced that the planned Fall workshop would be pushed into winter (Nov/Dec) due to time constraints. The original plan was to hold the workshop in Atlanta, Georgia, USA, however the floor was opened for other suggestions. James Johnston of U.S. Fish & Wildlife suggested that it be held in conjunction with the ERIM meeting scheduled in January of 1994. Mr. Johnston also agreed to help with the organization on that end.

Developing Nations Funding - The discussion of the workshop prompted a discussion of how members from developing nations could contribute considering the shortage of resources there (in particular, travel funds). The possibility of further participation through the use of e-mail and internet connection. It was agreed that this was, indeed, a possibility but there were still a number of countries that did not have internet connection. The possibility of obtaining funding from other, outside sources such as GTZ, UNESCO and other foreign aid agencies. Jeff Star (UCSB) volunteered to work with Dr. Ehlers on pursuing channels open to them.

Newsletters sent during 1993 - Three informational newsletters were sent to working group members over the course of the year. These mailings included information on upcoming Commission II events as well as WG II/2 events. The working group members were also provided information about the structure of the ISPRS Commissions and the names and contact persons for all of the Commission II Working Groups. All newsletters contained an updated member list. Whenever possible, the newsletters were sent via electronic mail.

Workshop Planning - Completed planning for and organization of upcoming WG II/2 workshop in New Orleans, Louisiana (see below in news). Successfully raised sufficient funding to provide travel support for five presenters from developing and Eastern European nations.

Working Group II/2 News

- Workshop on the "Requirements for Integrated Geographic Information Systems (IGIS)" will be held in conjunction with the ERA *Second Thematic Conference: Remote Sensing for Marine and Coastal Environments* in New Orleans, Louisiana, USA (2—3 February 1994).

- Working Group meeting to be held at the Commission II Symposium in Ottawa, Canada (June, 1994).

WG II/3 - "Technologies for Handling Large Volumes of Spatial Data"

by Chairman:	Dr. Ekow Otoo (Canada)
Co-Chairman:	Terry Fisher (Canada)
Secretary:	Cherian Chaly (Canada)
WG Members:	16

Terms of Reference

- raster data formats composition and transmission
- vector data format and topological models
- spatial data models
- database management systems for spatial data in a heterogeneous environment
- high volume storage media
- high performance I/O subsystems
- ISO 9000 and its interaction
- spatial data infrastructure

Accomplishments of Working Group II/3

During the past year, the working group sought and encouraged the participation of its members. Their interests have given way to a diversity of activities for the

Working Group to be involved in. Major technical and scientific activities of the past year include:

Examination of a number of different techniques for modeling spatial data, i.e., DIGEST, SAIF, ARC/INFO/VPF, SDTS, etc. Work will continue in this direction to develop a comparative study document of the different models of vector spatial data.

In March of 1993, presentation of a paper at the Fifth Canadian Conference on GIS, entitled *"Development of a framework for interoperability of a spatial database"*. A follow-up paper will be presented at the upcoming ISPRS Commission II Symposium in Ottawa, Canada

Working Group II/3 held a business meeting during the Fifth Canadian Conference on GIS, held in Ottawa, Canada from March 27 - 29, 1993.

The Working Group also organized and conducted a workshop prior to the opening of the Fifth Canadian Conference on GIS. The half day workshop on Object Oriented Spatial Databases was held on March 20, 1993, providing presentations and demonstrations.

Working Group II/3 held a business meeting at the Symposium of Spatial Data Handling, held in Singapore, during August of 1993. Among a number of items, WG members discussed the current issues of handling large volumes of spatial data and made plans for the Commission II Symposium in Ottawa from June 6 - 10, 1994.

A discussion meeting on methods of accessing Global Directories on Geophysical Data-sets, was held in September, 1993 in Ottawa, Canada.

Catalogue Interoperability Workshop, April 1993

State of Science and Technology Working Group II/3

The technology for the maintenance and processing large volume of spatial data continues to be a challenging problem in computing sciences and systems engineering. These problems are not restricted to spatial data alone, but span all areas where a high volume of data is captured and requires processing. Current database technology are very limited in coping with the size of the data volumes being considered here.

The problems are compounded by the fact that most spatial data resources are legacy data, they have a time dimension which is very essential to maintain for subsequent processing. The mere size of these datasets can not be handled by traditional database management systems, some records such as single images may span megabytes of disk space, and duplicating the data set for reliability purposes is very expensive. So far, the emergent technologies that have been used include: optical disk storage system, high density magnetic tapes and digital audio tapes, tape vaults and RAID technology and disk stripping techniques. These are technologies that although commercially available are still undergoing research.

The primary focus of Working Group II/3 is addressing these problems related to the storage and retrieval of very large spatial data and the technology for processing, and dissemination of large volumes of spatial information. In this regard, the Working Group's first year focused on identifying research projects, research centres and personnel around the world involved in various research activities related to the acquisition storage and processing of very large volume of spatial data. In the next two years, the activities of the working group will be split its responsibilities into three phases, as follows:

1) Problems of Spatial Information Representation and Models;

2) Methods and Technology for Data Storage, Access, Exchange and Exchange Standards;

3) Methods and Technology for High Speed Processing of Spatial Information Processing.

There are no distinct lines of separation of these three categories and hence phases of our work. For example, some of the issues may span the two phases of our work. These separations only suggest the broad areas of emphasis of the topics. In another sense our work may relate to those of other working groups as well. Problems of ***spatial information representation and models*** will deal mainly with methods and technology for the consistent, accurate and reliable representation and storage of spatial information. Some of the related issues that are expected to be covered in our working sessions are:

i. How should information of real world spatial objects be modeled and represented?

ii. Do current database management systems satisfy the large volumes of spatial data with respect to storage management, data access, and development of applications?

iii. What global index methods are there and if not what research initiatives are underway ?

iv. Can spatial information be maintained using database management system so that the traditional concurrency, backup and recovery mechanism used to institute satisfactory reliability of the data during usage and access.

v. How should temporal data be handled using current database management system or any recent advanced database management systems.

vi. What data compression techniques are most suitable for the economic storage of data without undue information loss?

vii. What meta-data representation techniques are there?

viii. What inter-relationships of various global directories, catalogues and services exist and how can this be effectively utilized?

The second phase of the work of the working group will address **methods and technology for data storage, access, exchange and exchange standards.** In addressing the problems of the storage of large volumes of spatial data, we distinguish between real-time and non real-time data acquisition, storage and access. The former relates to data acquisition where the time to store the data must meet some deadline. Namely that the time between successive storage of input must be less than the inter-arrival time of the input data. This is scenario is exemplified by the observation devices that capture and transmit observed data. However, a real-time data storage system must be able to organize and store the data at a faster rate than the rate of data acquisition. The non-real time scenario is exemplified by storage of map data. Here, the data such is captured, processed and to meet some standard quality control criteria after which it is submitted for storage. Both requirement pose some difficult problems yet to be answered by technological breakthroughs. The main issues of concern during this phase of our working activities are:

i. What real-time data compressing and decompression techniques are most suitable?

ii. Are there suitable database programming languages for communicating user accesses between application and stored data ?

iii. What spatial indexing methods can meet the high speed access selective access of data?

iv. What enhanced available technologies are there for large capacity data storage, beyond Optical Disk Storage, Digital Audio Tapes, Disk and Tape Vaults and RAID technology?

v. Is data stripping a satisfactory solution to high speed data input and output to high speed processing systems for spatial data or is the counterpart of large partitioned data units a rather satisfactory solution?

vi. Given the size of data typically retained, it is not feasible to conceive of backup and recovery strategies typical of traditional; transactional systems. Are there feasible strategies to backup and recovery other than replications ?

vii. Spatial data have been acquired and continue to be acquired and stored in a number of different formats. What common protocols can emerge as standards for interoperability in such heterogeneous databases and directory services?

The final phase of the working group's activities will be focused on technology for high speed processing in ***Spatial Information Systems***. Some of the major issues to be addressed will include:

i. Computing Environment for High Speed Processing of Spatial Data.

ii. Parallel processing strategies and Models of Computation.

iii. Technologies that address I/O bottlenecks.

iv. Real-time processing of spatial information.

v. Accessing and processing techniques for temporal data.

vi. Constraints involved in the modeling of geophysical phenomenon and processes.

Most of these issues require high computer intensive resources and more so in the area of preventing the processing action from data starvation due to I/O bottlenecks. We will examine solutions that coordinate the computational algorithms with the input output operations to achieve very high through puts. However, these issues pose challenging algorithmic problems.

Working Group II/3 News

The group will continue to cooperate with personnel from the industry, government institutions and academic research institutions to study and report on interesting findings. In particular, efforts will be made by members of the Working Group to participate in conferences and workshops where discussions focus on such technologies. The working group will meet at least once a year to discuss and report on individual activities during the year. This discussion would culminate in a report to be maintained as part of their yearly activities. The group will maintain close liaison with international standards organizations, such as the Committee for Earth Observation Satellites (CEOS), the Consultative Committee for Space Data Systems (CCSDS) and the NASA Catalog Interoperability Working Group. Interesting developments, research results and findings will be reported in a series of meetings leading to the ISPRS Congress in 1996. Proposed meeting times should be consistent with some scheduled conferences and workshops such as:

Presentation of a follow-up paper of ISPRS Commission II Symposium to be held in Ottawa, Canada from June 6 - 10, 1994.

Organizing and presenting a workshop prior to the opening of the Commission II Symposium in Ottawa, Canada in June of 1994.

Business Meeting to be held at the International Symposium on Data Handling, Edinburgh, Scotland.

Attendance at the ASPRS/ACM Annual Convention in the United States.

WG II/4 - "Systems for the Processing of Radar Data"

by Chairman: Dr. Robert O'Neil (Canada)
Co-Chairman: Dr. Hiroshi Kimura (Japan)
Secretary: Dr. Marc D'Iorio (Canada)
WG Members: 46

Terms of Reference

- development of methodologies to process, analyze and interpret Synthetic Aperture Radar (SAR) imagery.

- validation of applications of imagery acquired by space-borne instruments such as ALMAZ, ERS-1, JERS-1, RAOARSAT, ERS-2, and ENVSAT

- evaluation of SAR interferometric techniques for elevation determination.

- prepare and publish sample data set with various scene types for the evaluation of SAR analysis algorithms.

Working Group II/4 Accomplishments

May 1993 - working group assembled by invitations sent to international community active in SAR data analysis field.

August 1993 - informal meeting held of members attending the IGRSS'93 Tokyo, Japan.

October 1993 - solicited abstracts and planned sessions on SAR data analysis at the ISPRS Commission II Symposium during June, 1994 in Ottawa, Canada.

Working Group II/4 News

- Definition of the CD-ROM data set is underway. Research groups world wide will be invited to contribute sample digital SAR scenes acquired with various SAR instruments and processors.

- A meeting of the Working Group is being planned to coincide with the ISPRS Commission II Symposium June 1994 in Ottawa.

- Several groups have been contacted to determine willingness to undertake or contribute to a workshop on SAR data analysis at the ISPRS Congress July 1996 in Vienna, Austria.

WG II/5 - "Integrated Production Systems"

by Chairman:	Dr. Atef A. Elassal (USA)
Co-Chairman:	Dr. M.M. Radwan (Netherlands)
Secretary:	Dr. Roop C. Malhotra (USA)
WG Members:	15

Terms of Reference

- the role of GIS, image analysis systems (IAS), remote sensing and photogrammetric technologies in the design of Integrated Production Systems (IPS)
- design and development aspects of IPS, including benchmark tests for system evaluation
- standards/methodologies for data communication, data exchange, operational data flow and output for IPS
- development of total digital platforms and spatial information infrastructure for the management of heterogeneous corporate databases

Working Group II/5 Accomplishments

July 9, 1993 - letters of invitations to join the WG were sent world-wide to persons in academia, industry and government agencies.

Working Group Membership - there are now fourteen members, including two reporters, representing ten countries.

Circular letter - the task of collecting information with the objective of publishing a user guide for Integrated Production Systems: "IPS User Guide "was identified by WG chair. Letters to the members were sent out to solicit information on existing IPS world-wide.

Circular letter - a sample of an IPS was prepared and enclosed in a circular letter to the membership for information content and template in which the information may be presented.

November 8, 1993 - a working group meeting was at NOAA, Silver Spring, Maryland immediately following the 1993 ASPRS Minnesota, Convention.

State of Science and Technology of Working Group II/5 Topics

Following the successful Working Group Meeting in November, and number of issues concerning Integrated Production Systems rose to the forefront. Highlights of the WG Meeting and these issues are as follows:

- it was recognized that the spatial data information collected by photogrammetric based IPS must be linked to some GIS for analysis and end products.
- for wider use of IPS, existing IPS must be characterized for functionality, hardware/software, and other relevant information to assist users in setting up IPS for data collection to cater to the community needs. This boiled down to the WG task identified by Dr. Elassal and Dr. Malhotra in the initial search for WG tasks.
- it was suggested that to better understand the scope of the WG and the IPS, the systems that the WG addresses must be called: "Integrated

Photogrammetric (including Remote Sensing) Data Acquisition Production System". In other words, IPS must be photogrammetric based systems with data acquisition functionality, leaving the task of data analysis to some sort of GIS.

- Thus, the IPS must have the following functionalities:
  - Spatial data collection, editing and storage
  - Data transfer
  - Integration of data sets (vector/raster, from different sensors)
  - User interfaces to GIS etc.

- The word integration used in IPS may be restricted to "Integration of various processes for spatial data acquisition", such as: triangulation, compilation. Data analysis is relented to GIS. Linkages between IPS and GIS may he considered as a part of IPS.

- Input to IPS user guide should include:
  - description of all the INPUT data to the IFS
  - description of data flow or processes in IFS
  - description of functionalities of IFS
  - definition of accuracy of the data set and of the OUTPUT data

Working Group II/5 News

- IPS User Guide Planned Milestones:

March, 1994: first cut at the template for reporting an IPS

June, 1994: interim report (include at least one IPS from industry)

1996 Congress: final IPS user guide an current IPS.

- Implementation action plan for IPS user guide:

Nov./Dec. 1994: outline and example from the Chair to be sent to WG membership with inputs from members and finalize template

March, 1994: Solicit inputs from other agencies

June, 1994: Interim IPS user guide.

1996 Congress: Final IPS user guide an current IPS

Inter-Commission WG II/III - "Digital Photogrammetric Systems"

by Chairman:	Prof. Ian Dowman (U.K.)
Co-Chairman:	Dr. David McKeown (USA)
Secretary:	Zubbi Nwosu (Switzerland)
WG Members:	186

Terms of Reference

- digital photogrammetric workstation design

- design issues for computational tasks using parallel processing, multi-processors and task specific architectures

- visualization techniques including stereo, spatial, temporal, and animation issues in computer graphics

- multimedia techniques for imagery, scanned documents, sound and video

- automated cartographic compilation systems (algorithms and system interfaces)

- integration of photogrammetric techniques and models into computer vision systems for the analysis of remotely sensed imagery
- integration with input and visualization devices such as image scanners and raster plotters
- human-computer (human-machine) interface issues in photogrammetric system design
- integration with spatial databases including digital maps, terrain models and remotely sensed imagery

InterCommission WG II/III Accomplishments

The Working Group Activities during the past year have centered on preparing the program for the period up to the Congress in 1996. A Newsletter was distributed in March based on the mailing list from 1988-92 and as a result of that a revised list has been prepared.

In April 1993a WG business meeting was held in Orlando, Florida on the occasion of the SPIE International Symposium on Optical Engineering and Photonics in Aerospace and Remote Sensing which included a conference on Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision co-chaired by Dave McKeown (co-chair of ICWG II/III). see below for more details. The discussion at the business meeting started on the topic of systems but quickly moved on to standards and bench marks for digital systems. There was clearly a strong feeling that users needed some guidance on the performance to be expected from a system. It was noted that this was not an easy task because of the wide range of systems offered, from the Intergraph Image Station to the DMS. There was also a feeling that such tasks can take a considerable amount of time with little intellectual satisfaction.

There was general agreement that the terms of reference were broad but appropriate and it was the breadth which distinguished the working group from other more specialist groups, especially those working with algorithm development. The view was put that the group should look to the end user and that there was a trend in the USA towards the use of orthoimages and that he user might well do his own revision using the orthoimage.

The Working Group is considering its position on standards and is planning to hold a session at the Commission II Symposium on the subject.

Since then the chairperson has had discussions with chairmen of other working groups about joint meetings and collaboration and also discussed the possibility of holding a working group meeting in South East Asia, with the head of the Remote Sensing Centre at the Technical University of Malaysia.

The state of science and technology for InterCommission WG II/III

There have been a few meetings since August 1992 which have included sessions on DPWs. The most notable have been:

SPIE "Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision", in Orlando, U.S.A. This meeting stressed the importance of both close range and topographic applications of digital systems. A review paper by Heipke set out the current status of digital systems. The conference attracted good quality speakers and a large audience and good discussion took place between the photogrammetric community and the machine vision community. A full report of the meeting is included in The ISPRS Journal of Photogrammetry and Remote Sensing Vol. 48 No 5, Oct. 1993.

Photogrammetric Week, Stuttgart, Germany

The Photogrammetric Week concentrated on digital photogrammetry and was the occasion of the launch of the Zeiss PHODIS ST for three dimensional plotting. Papers also showed that digital orthophotographs are becoming increasingly used by production organizations for a number of applications.

There is a steady increase in the development and use of DPWs. Apart from the Zeiss PHODIS, the development of DFWs for MOMs data is reported and a workstation has been developed in India at the Advanced Data Processing Research Institute, Department of Space. There is also increasing use of integration of data from different sensors, ISTAR have worked with SAR and SPOT for example.

Another area of interest to the Working Group is multi media systems. SUN have announced a range of SPARC1O multi media workstations which include cameras and real time video capture/compression cards.

InterCommission WG II/III News

- June 1994, Commission II Symposium, Ottawa, Canada
- September 26 - 28th, 1994, Commission III Symposium
- (Date not yet confirmed) Workshop and seminar in Malaysia
- Spring 1995 Conference in USA
- August 30th - Sept. 1st, 1995 Joint meeting in Stockholm with WG 111/2
- July 1996 Congress

Commission II Special Topic WG - "Upgrading Photogrammetric Instruments"
by Chairman: Dr. Klaus Szangolies (Germany)

Terms of Reference

- survey recent and future developments in analog and analytical photogrammetric systems and identify/classify the various types of add-on devices and/or instrument modification
- evaluate the impact of upgrading photogrammetric systems with digital devices and computer technologies on the instrument productivity, and its life-span and users

Accomplishments of Commission II Special Topic

Circular letters to manufacturers of photogrammetric equipment and national reporters of the Special Group

Liaison and discussion with manufacturers during:

May 5 - 7, 1993 - Geotechnica, Cologne, Germany

May 25 - 28, 1993 - Quality Stuttgart, Germany

September 16 - 17, 1993 - Deutscher Geodatentag, Augsburg, Germany

September 27 - October 1, 1993 - attendance at "International Mapping from Space", Hannover, Germany.

Invited 6 persons to prepare papers for presentation at Commission II's Symposium in June, 1994 in Ottawa, Canada.

State of Science and Technology of Commission II Special Topic:

The stereoplotter for the graphical 3-D representation of data extracted from stereo photographs is still the most widely used instrument in both topographic map production and industrial and architectural photogrammetry.

The number of stereoplotters employed in the world today is estimated at 5000. They were manufactured between 1950 and 1990. Stereoplotters were built with quite different designs. There are optical, mechanical, and analytical models with approximation solution and models with mathematically exact solution, first, second and third order models, i.e. expensive models featuring high accuracy and universality and simple, small budget priced models. The purchase price ranged between DM 50,000 and 500,000.

Since many of these expensive machines are far from worn out, there is a great demand for:

- modernization
- complementation by digital output and input modules
- linkage to computers for automation and increases in productivity.

Special Topic News

- Business Meeting May 31 - June 3, 1994 during Symposium of Commission IV in Athens, Georgia, U.S.A.

AUTHORS and COAUTHORS INDEX
ISPRS COMMISSION II