# Computer-Based Aerial Image Understanding: A Review and Assessment of its Application to Planimetric Information Extraction from Very High Resolution Satellite Images

by Bert Guindon

## *RÉSUMÉ*

*Au cours des cinq prochaines années, une série de capteurs satellitaires seront lancés offrant la possibilité d'acquérir des images monochromes à des résolutions spatiales variant de 1 à 3 mètres et des images multispectrales à des résolutions plus faibles variant de 4 à 15 mètres. La cartographie topographique (Konecny, 1996) à grande échelle (1:50000 ou plus) constituera une application potentielle importante pour ces données. Des techniques d'extraction d'information automatisées seront donc requises pour permettre l'exploitation de cette source de données à grand volume. Quoique des progrès importants aient été réalisés dans des secteurs comme la génération de modèles numériques d'élévation (MNE), d'autres secteurs, notamment l'extraction des caractéristiques plan-imétriques, font encore appel à une main-d'oeuvre intensive (Leberl, 1994). La communauté civile de télédétection qui a surtout traité des images multispectrales à résolution moyenne ou faible (10 à 1000 mètres), a concentré ses efforts sur l'exploitation des attributs spectraux plutôt que les attributs spatiaux spécifiques des images pour déterminer l'utilisation thématique du sol. Quoique ces réalisations aient été significatives pour les besoins de la cartographie synoptique, elles seront d'intérêt limité pour la cartographie à grande échelle dans le contexte des nouveaux capteurs. D'autre part, des recherches ont été réal-isées parallèlement au cours des deux dernières décennies pour développer des systèmes informatisés d'analyse d'images aéroportées. Ces recherches ont été financées en grande partie par les militaires, ce qui explique qu'on ait négligé les aspects reliés au domaine des applications. Toutefois, cela a permis le développement d'une série de technologies d'analyse spatiale conçues pour des images monochromes à haute résolution spatiale qui peuvent, en théorie, être généralisées à des applications portant sur une variété plus grande de scènes. Dans cet article, on présente une revue des derniers développements dans le domaine des systèmes d'analyse d'images, incluant un survol des méthodologies principales, de même qu'un inventaire des systèmes intégrés d'analyse d'images les plus importants. En conclusion, nous présentons une discussion sommaire sur les secteurs spécifiques vers lesquels il faudra orienter le développement dans le futur dans le contexte de l'application de cette technologie au domaine de l'automatisation de l'extraction d'information planimétrique.*

## *SUMMARY*

*Within the next five years a series of satellite sensors will be launched that will be capable of providing monochrome imagery with spatial resolutions in the range 1 to 3 metres and multispectral imagery at lower resolutions in the range of 4 to 15 metres. A key potential application of such data will be fine scale (i.e. 1:50000 or better) topographic mapping (Konecny, 1996). To exploit this high volume data source, automated information extraction techniques will be required. While great strides have been made in areas such as digital elevation model (DEM) generation, other areas, in particular planimetric feature extraction, remain operator intensive (Leberl, 1994). The civilian remote sensing community, which has dealt largely with moderate to low resolution (10 to 1000 meter) multispectral imagery, has concentrated on exploiting spectral rather than detailed spatial image attributes to infer thematic landcover. While these achievements have been impressive for synoptic mapping purposes, they will be of limited value for fine scale mapping from the new sensors. On the other hand, parallel research has been on-going for the past two decades to develop computer-based aerial image under-standing (IU) systems. This research has been largely military-sponsored and as a result has had a narrow applications focus; however, it has resulted in the development of a host of spatial reasoning technologies which are tailored to high spatial resolution monochrome images and which can, in theory, be generalized to apply to a broader range of landscapes. In this paper a review is presented of the state of-the-art in aerial IU, including overviews of some of the key methodologies, as well as a survey of major integrated IU systems. The paper concludes with a summary discussion of selected areas requiring further development if this technology is to be applied to automating planimetric information extraction.*

---

- B. Guindon is with the Canada Centre for Remote Sensing, 588 Booth Street, Ottawa, Canada, K1A 0Y7

## INTRODUCTION

Most image processing algorithms developed and employed by the civilian remote sensing community have been tailored for the analysis of medium to low spatial resolution (10-1000 meter) images. Typically these algorithms exhibit the following characteristics:

(a) Processing is pixel-based, relying on spectral dimensionality rather than spatial context to accurately 'classify' images.
(b) When spatial information is included, it is based on the relative radiometric attributes of pixels within fixed windows, usually referred to as image texture (e.g. Haralick, 1979). Generally, textures are encoded in the form of pseudo-bands which can be readily processed in conjunction with conventional spectral bands using per-pixel processes such as supervised classification. Empirical post-classification processing is also used to encapsulate context, for example, the application of classification map filtering.

As image spatial resolution increases, greater local radiometric heterogeneity can be expected and hence the need for more sophisticated methods to directly quantify and interpret spatial information. As resolutions approach 1 metre, many man-made objects such as buildings will be discernable as discrete image features. As a result, traditional low resolution thematic classes such as 'suburban' will disaggregate into a diversity of functionally-related subcomponents (e.g. house, driveway, lawn, street).

In parallel, but in general isolated from civilian remote sensing activities, there has been on-going computer vision research directed at the analysis of high resolution monochromatic aerial photos. These endeavours have led to the development of a rather different approach to scene interpretation, usually referred to as 'image understanding' (IU). The key characteristics of IU are summarized below.

(a) IU involves feature, rather than per-pixel, based processing and manipulation. Images are partitioned into 'feature' components such as segments, lines, etc. Class labels are assigned to each such 'feature' based on its intrinsic attributes or its context vis-a-vis neighbouring features.
(b) While per-pixel classification relies on 'training' pixels to quantify statistical image attributes of classes, image understanding relies on knowledge of the generic structural characteristics of physical objects (e.g. buildings, roads, etc.), their expected spatial and functional relationships and imaging models to predict their corresponding image feature characteristics.
(c) In the most general case, interpretation is conducted at a range of levels of spatial context, employing both image-directed (bottom-up) and model-driven (top-down) processing. For example, initial labels may be assigned to features based on their in-situ attributes. This is followed by model-driven spatial reasoning to resolve conflicting classifications, infer missing object components and construct scene models of increasing spatial extent.
(d) Since military applications have been a prime driver of IU technology, limited-goal strategies have been emphasized. In such scenarios, scene interpretation is limited to the location and identification of only a subset of 'objects', usually man-made ones such as aircraft, vehicles, special-purpose buildings, etc. As a result, a typical processing strategy may involve two steps, first, full scene 'reconnaissance' processing to locate regions of interest (e.g. an airfield) followed by in-depth analysis of these restricted regions to detect and identify those specific objects of interest (e.g. aircraft).
(e) Early work in IU was directed at the analysis of single images hence research addressed problems related to the derivation and interpretation of monocular cues. For example, attributes related to height had to be inferred from the presence or absence of shadow. Recent work has addressed the analysis of digital stereo image pairs and the integration of photogrammetric methods with IU image processing.
(f) Great effort has been put into both using existing collateral information in scene interpretation and generating new collateral information from interpreted images. These goals are dictated by the nature of the major applications of most military-sponsored IU systems, which are reconnaissance and intelligence gathering. In a typical scenario, the system operator (photointerpreter) begins with a 'site folder', i.e. a database which includes both a current site model (i.e. a summary of the site contents, e.g. locations and types of buildings, roads, airport runways, etc.) and the data from which that model has been derived (e.g. raw and interpreted images). The model itself may consist of both graphical (i.e. digital maps, images, etc.) and textual data (i.e. a summary description). The 'ideal' IU system goals are to (a) ingest new imagery, (b) spatially integrate it with existing data, (c) derive a scene model from it using as a seed the existing site model, (d) report the locations of potential change and the nature of that change and (e) aid the operator in assessing and validating changes and updating the site model (including the generation of textual reports). The integration of image and textual data is not one which has been seriously addressed in civil remote sensing but which is of pivotal importance in, for example, automated map updating.

Given the approaching launch of operational civilian sensors with geometric fidelity rivalling that of aerial photography, it is imperative that existing IU tools be critically assessed within the context of specific civilian remote sensing applications. In this paper we survey IU research and discuss its potential for automating planimetric feature extraction in support of topographic mapping.

# REVIEW OF RELEVANT WORK IN AERIAL IMAGE UNDERSTANDING

There are numerous textbooks which describe fundemental image processing techniques used in IU as well as knowledge issues (e.g. Ballard and Brown, 1982; Haralick and Shapiro, 1992; Sonka *et al*., 1993). The reader is referred to these for background material. As this paper presents an in-depth review of selected topics relevant to planimetric feature extraction, we will refer directly to specific research articles. Two key ingredients of all IU systems are (a) baseline image processing functions for the extraction and description of primitive image features (i.e. edge and segment structures) and (b) 'knowledge engineering' issues which encompass spatial reasoning strategies and methods for knowledge encapsulation. We briefly review these elements in this section. This is followed by a survey of current IU system implementations.

## Image Feature Extraction and Description

### *Region Boundary Extraction*

Boundary delineation is a complex process involving (a) estimation of edge magnitude and direction at every pixel location, (b) edge thinning, (c) thesholding to eliminate noise-related responses, (d) hierarchical linking of edge information first at the pixel level into logically-connected chains and ultimately into groupings of chains, (e) gap filling and (f) detection of nodes, i.e. points such as corners which link intersecting edge chains. In general only a single edge type is being sought, however, there are numerous formulations of detection criteria, each of which lead to a distinct set of edge response templates. The most frequently employed algorithms in aerial image processing include (a) an ideal zero-width step (Nevatia and Babu, 1980), (b) derivative of a gaussian (Canny, 1986) which rationalizes 3 measures of performance, namely 'good' detection (low probability of missing real edges and detecting noise) , 'good' position estimation and single response to each edge, which also accomodates edges of a finite width, and finally (c) the Laplacian of a Gaussian or second derivative of a gaussian (Marr and Hildreth, 1980).

Nevatia and Babu also propose methods for the remaining steps although linking is limited to the labelling of linear chains and the pairing of antiparallel chains for the purpose of detecting roads. Important refinements and extensions applied in aerial image IU systems include utilization of B-splines to model chain curvature thereby aiding in corner detection (Medioni and Yasumoto, 1987), incorporation of extrapolation methods in road tracking (McKeown and Denlinger, 1988), linking of linear chains to locate shadow cues in the identification of rectangular buildings (Huertas and Nevatia, 1988), linking collinear line fragments (Venkateswar and Chellappa, 1992), perceptual grouping methods to identify collated features (Mohan and Nevatia, 1988, 1989) and 'model-driven' edge detection which utilizes cost functions to derive best fitting curvilinear features through noisy edge data (Fua and Hanson, 1988).

### *Segmentation*

Segmentation can be defined as the process of subdividing an image into a set of 'homogeneous' regions. In the simplest case, the homogeneity criterion can be equated to uniform image brightness, however, in the current application this must be generalized to include uniform texture as well. Segmentation techniques can be conveniently classed as boundary-based, region-based or hybrid (combination of the two). As with edge detection, the literature on this topic is too vast to cover in detail (e.g. see for examples recent surveys by Pal and Pal (1993), Reed and du Buf (1993)), so we limit this discussion to papers dealing with techniques which have been directly applied to aerial images.

Boundary-based methods involve tracing closed boundaries around regions, usually via edge detection and linking. Perkins (1980) proposed an effective method, involving only local operations which addresses two practical issues, namely, gap filling and the discrimination of edge responses of region boundaries from those associated with intra-region texture. Gaps are filled through a succession of edge dilation/thinning operations while connectivity criteria are employed to distinguish between texture and boundary-induced edge responses.

Region growing, on the other hand, involves seeding a large number of regions then, through a series of split-and-merge operations based on region homogeneity and adjacent-region similarity criteria, attaining a final consistent partitioning. Recently, Pan (1994) proposed a promising variation which uses iterative smoothing to initially fragment an image into a large number of regions. The subsequent merging is governed by a goal of global (i.e. scene-wide) information preservation, a concept proposed and developed by Leclerc (1988, 1989).

Both boundary and region methods each have their limitations, hence hybrid methods have also been proposed. In general these have been developed for restricted object types and landscapes (e.g. road finding (McKeown and Denlinger, 1988) and large scale building delineation (Liow and Pavlidis, 1990)). In addition, they tend to go beyond simple image processing and incorporate object knowledge to guide the overall process. For example Liow and Pavlidis look for cast shadows to find high contrast edges which are most likely associated with building roofs, then use region growing to find the remaining sun-facing sides.

## Knowledge Engineering Issues

In this subsection we address two key issues related to the utilization of computer-based reasoning namely process control strategies and knowledge representation. More details can be found in papers addressing the role of artificial intelligence in photogrammetry (Sarjakoski,1988; Forstner, 1993) and system implementation issues ( e.g. Matsuyama, 1987; Xu, 1992).

### *Process Control*

Machine interpretation of complex imagery requires a comprehensive suite of processing modules each with unique inputs, functional limitations and outputs. Process selection and control is then a critical issue and, at its highest level, is a decision as to whether processing should be image or object-driven. For a excellent overview discussion of this topic see Matsuyama (1987). In an image-driven scenario, also referred to as bottom-up control, processing proceeds from raster image to image feature extraction to feature description to recognition. Bottom-up processing is the approach taken in conventional image processing. On the other hand, model-driven or top-down control begins with a set of assumptions/properties of an expected object list. These object properties are tested against a sequence of image representations of ever decreasing levels of abstraction. Through a series of test/validate processes one eventually reaches an interpretation at the parent image level. The two control strategies each have their own inherent advantages. For example, a bottom-up approach is particularly useful for providing initial estimates of labels for a subset of features, while a model-driven approach does well at inferring 'missing parts' given an initial incomplete interpretation. Effective system performance requires a combination of the two within the confines of an iterative refinement strategy (Nicolin and Gabler, 1987; Matsuyama, 1987; McKeown *et al*., 1989).

### *Knowledge Representation*

IU systems typically contain a diverse set of knowledge whose elements can be conveniently grouped into object and image-related categories. Object-related elements can be further categorized as one of three forms, (a) intrinsic properties such as object size, shape, etc. (see for example Perkins *et al*., 1985), (b) relational information which defines the expected spatial relationships among collections of objects of different classes (e.g. houses are spatially adjacent to driveways) and (c) procedural knowledge, typically in the form of 'condition-action' pairs which dictate actions to be undertaken to test a labelling hypothesis. Symbolic representations of objects, such as in the form of generalized cylinders, has proven to be successful (Bogdanowicz and Newman, 1989; Brooks, 1981) but have limited applicability beyond a restricted range of man-made targets such as aircraft components. On the other hand, there are a number of knowledge characteristics which a respresentation model should be capable of handling. First, object knowledge should be expressable in a generic or proto-typical form, making it applicable to a wide range of instances (i.e. landscapes). Second, some object knowledge is strictly semantic, which makes its definition difficult to formulate within a precise mathematical framework. And third, for most if not all physical object classes of interest there will be no single definitive attribute that will unambiguously identify them in an image. As a result, the knowledge model must allow for cumulative evidence gathering to support intepretation hypotheses, imprecision in evidence and multiple interpretations of individual image features. Frame-based representations provide a viable means of encapsulating these properties (Mundy *et al*., 1992) while object-oriented programming languages provide a suitable implementation environment.

The effectiveness of image processing (IP) tasks in extracting features which can be easily related to physical objects of interest requires expertise in both the selection of processing parameters and in sequencing task executions. While most commercial IP packages rely solely on human expertise, pioneering work in developing rule-based systems to exploit this knowledge has already been carried out both in the remote sensing community (e.g. Goodenough *et al*., 1987; Fung *et al*., 1993) and in the IU community (e.g. Matsuyama, 1987; Bjorklund *et al*., 1989). In many IU scenarios, rule selection is dictated by a limited interpretation goal, e.g. the identification of examples of a single class of object (e.g. houses) based on expected in-situ image feature properties (e.g. rectangular in shape and with a specific areal extent). In general achievement of this goal will be at the expense of overall scene processing performance. A frame-based encapsulation can also be employed both to define required in-situ image feature properties and the procedures (IP tasks) necessary to extract them.

Finally, independent system developments at a variety of academic and private sector institutions have led both to duplication of effort and difficulties in results sharing (McConnell and Lawton, 1988). As a result, there are on-going efforts to develop a common software development environment (Mundy *et al*., 1992; Bremner *et al*., 1996) and to integrate key IU components on a single platform (Edwards *et al*., 1992).

## REVIEW OF MAJOR IMAGE UNDERSTANDING SYSTEMS

**Table 1** summarizes some of the major systems developed to date for aerial IU. Each can be categorized into one of three classes, (a) systems which attempt to provide full scene interpretation, (b) goal-directed systems which attempt to delineate and characterize a limited set of objects (e.g. buildings, roads, vehicles) and (c) 'interpreter-aid' systems which provide automated functionality in areas such as image rectification, diverse data integration and enhanced visualization but which still rely on human interpretation. Rather than presenting overviews of all of the systems in **Table 1**, a representative subset is discussed in detail. This is supplemented with additional information of unique features of the remainder.

### SPAM (System for Photo Interpretation of Airports Using MAPS)

SPAM was first developed at the Carnegie-Mellon University in the mid-1980's to analyze airport scenes (McKeown *et al.*, 1985) but has undergone numerous improvements and has been subsequently applied to urban scene interpretation(see McKeown *et al*., 1989, Harvey *et al*., 1992, McKeown *et al.*, 1994). It is of particular interest because of its spatial reasoning methodology and its comprehensive functionality.Airports are ideal candidates for study because of their well defined, generic functional components and component spatial relationships. SPAM begins by segmenting an image into a set of regions and deriving a set of spatial/spectral attributes for each region. Classification proceeds through a bottom-up hierarchy of interpretation phases. Phase 1 consists of generating a classification of as many regions as possible based solely on region attributes and knowledge about the scale and structure of class objects. These classified regions are referred to as fragments. Phase 2 uses knowledge of the spatial relationships among classes to check the consistency of adjacent fragment interpretations. In

**Table 1.**
**Summary of Aerial Image Understanding Systems.**

| System Name | Reference(s) | Comments |
|---|---|---|
| ACRONYM | Brooks (1981) | - symbolic model representation of features (generalized cylinders)<br>- bottom-up reasoning only<br>- applied to aircraft recognition |
| ARF<br>( A Road Tracker) | McKeown and Denlinger (1988) | - high resolution road tracker |
| BABE<br>(Built-up Area Building Extractor) | Shufelt and McKeown (1993) | - building detection based primarily on monocular cues (e.g. shadows)<br>- applied to large rectangular buildings in dense urban areas |
| CANC | Mohan and Nevatia (1989) | - employs perceptual grouping of low-level collated image features<br>- applied to delineation and matching of rectangular buildings in stereo image pairs |
| CYCLOPS | Barnard (1990) | - automated cartography including orthorectification, DEM reation, perspective viewing |
| COBIUS<br>( Constraint-Based Image Understanding System) | Bjorklund *et al*. (1989) | - key functions include generic domain object representations, knowledge control, compensation for unreliable segmentation<br>- applied to aircraft recognition |
| MULTIVIEW<br>(Multi-Image Building Extraction) | Roux *et al*. (1995) | - building detection and delineation based on direct matching of cues in multiple images |
| PACE | Corby *et al*. (1988) | -compiles multiple images into a common 3-D reference frame<br>- synthetic image generation/display<br>- applied to aircraft surveillance |
| ROADF | Zlotnick and Carnine (1993) | - road tracking system |
| SCORPIUS | Bogdanowicz and Newman (1989) | - object-directed recognition of man-made objects<br>- automated generation of assessment reports to aid human interpreters<br>- symbolic representation of objects (generalized cylinders) |
| SIGMA | Matsuyama (1987) | - both bottom-up, top-down processing<br>- frame-based knowledge representation<br>- applied to suburban scenes |
| SPAM<br>( System for Photo Interpretation of Airports Using MAPS) | Harvey *et al*. (1992)<br>McKeown *et al*. (1989)<br>McKeown *et al*. (1985) | - multiple spatial levels of reasoning<br>- accuracy assessment, performance evaluation modules<br>- automated knowledge acquisition<br>- applied to airport and suburban scenes |
| TRIPLE<br>(Target Recognition Incorporating Positive Learning Expertise) | Bhanu and Ming (1988) | - emphasis on system learning<br>- applied to military vehicle detection and identification |
| 3DIUS | Fischler and Bolles (1994) | - human interpreter aid<br>- major components include image integration via photogrammetric tools, texture-mapped rendering, object model-image integration<br>- builds upon Cartographic Modelling Environment (CME) of Hanson and Quam (1988) |

Phase 3 an attempt is made to group collections of local fragments into functional areas thereby creating a larger scale interpretation. In this process, knowledge of the the functional components and their spatial relationships can be used either to rationalize ambiguous fragment interpretations or to predict the classes of previously unclassified regions based on a 'missing part' analysis of the area. Finally, functional areas are merged and rationalized to generate a full scene model.

A significant effort has been put into the analysis of system characteristics and performance of SPAM. A complex user interface was developed (a) to employ 'ground truth' in the form of manual segmentations to update attribute characteristics for classes, (b) to edit/delete/modify rules and (c) track the performance impact of individual rules within the context of a complex knowledge environment. Assessment of scene models can be undertaken through comparison with human interpretations. Recent enhancements include methods for automated knowledge acquisition (Harvey *et al*. 1992), the employment of parallel processing to improve computational performance (McKeown *et al*., 1994) and tools for refining knowledge bases (Harvey and Tambe, 1993).

## BABE (Built-Up Area Building Extractor)

The previously described system, SPAM, was designed to generate full scene models, i.e. by labelling every image segment. Most IU systems have much more modest goals, either to detect and describe selected object classes or to aid a human operator in manual interpretation. BABE is an early example of the former class and is of particular interest because it's evolution parallels many of the major directions in aerial image understanding research.

The original BABE system addressed the problem of detecting building candidates in single, near-nadir viewing images. In general these scenes encompassed only dense urban settings such as the cores of large cities where man-made objects constitute most of the ground cover and buildings tend to be rectangular high-rises with flat roofs. Building detection is accomplished through a combination of edge detection and line-corner analyses to first identify closed rectangular image features (Huertas and Nevatia, 1988). Monocular cues such as shadow area delineation by grey level thresholding and detection of roof-shadow boundaries are then sought in order to distinguish raised from ground-level enclosures (Irwin and McKeown, 1989; Shufelt and McKeown, 1993; Lin and Nevatia, 1996).

In simple situations, for example isolated buildings on level ground, monocular shadow cues such as shadow length can be used to infer building height as well. In general, however, pairs of matched stereo images are required. Conventional area based image matching is of limited use to estimate the parallax shift due to height because of elevation discontinuites associated with building-ground level boundaries and structural occlusions associated with off-nadir viewing. Alternate approaches have been investigated including 'waveform' matching of radiometric profiles extracted along epipolar lines (Perlant and McKeown, 1990; Hsieh *et al*., 1992) combined with image segmentation to identify flat roofs of near constant reflectance (McKeown and Perlant, 1992). Ford and McKeown (1992) have also shown that if multispectral rather than monochrome imagery is available, improved recognition can be achieved from the addition of cues derived from conventional spectral classification.

Another important recent trend has been the incorporation of photogrammetric modelling within aerial IU systems. As a result, rigorous analysis can now be undertaken of images acquired at extreme oblique viewing angles. With the aid of 'vanishing point' or perspective geometry, edge detection can be improved since the absolute orientation of line segments in object space can be inferred. Intersecting vertical-horizontal line pairs (e.g. associated with building walls) can be distinguished from intersecting horizontal pairs (e.g. roof or ground level boundaries) thereby greatly aiding in the extraction of self-contained building structures (McGlone and Shufelt, 1993). In addition, the ability to accurately relate image and object space allows one to successfully execute feature matching in images acquired with very different viewing perspectives (Mohan and Nevatia, 1989; Murphy, 1993; Venkateswar and Chellappa, 1992; McKeown *et al*., 1994), locate buildings in images which match an object space template (Meuller and Olson, 1993; McGlone and Shufelt, 1994), and extract 3-D object space models of buildings from real scenes and subsequently view them from any arbitrary viewpoint (Roux and McKeown, 1994; Shi and Shibasaki, 1996).

Because of the difficulties in extracting coherent building structures and matching them in stereo imagery, there have been few studies which address critical accuracy issues such as horizontal positioning of buildings and their absolute height estimation (Dessard and Jamet, 1995).

## ARF (A Road Follower)

Image processing to extract road networks consists of three steps, (a) road finding, the process of locating 'seed' points on one or more of the roads present in a scene, (b) road tracking, which involves tracing a road from a starting seed location and (c) road linking, the connection of discrete road sections into a consistent, logical network (Zlotnick and Carnine, 1993). Extensive research has been undertaken to extract roads in low resolution satellite imagery (see, for example, Wang and Liu (1994) for a comprehensive review of this topic). In such imagery, roads widths are typically less than or comparable to the scale of the processing windows of feature operators (typically high pass filters) employed to highlight them and therefore roads can be considered to be 'line-like'. On the other hand, on metre-resolution imagery, roads are fully resolved and exhibit two distinct boundaries as well as internal structures and textures. This added complexity, while making road delineation potentially more complicated, also opens the door to the utilization of a wider suite of detection tools, such as structural analysis, and to the extraction of a more comprehensive set of attributes (e.g. width, surface material, traffic characteristics).

McKeown and Denlinger (1988) developed ARF to track roads from manually-located seeds. The approach involves extraction of a local image intensity profile, normal to the road direction at a seed location, then tracking of the road based on region and correlation-based profile matching. This differs from road boundary detection techniques which rely on edge following (e.g. Heipke *et al*., 1994). Additional reasoning methods are included to detect a suite of road characteristcs including intersections, road width changes, surface changes, overpasses, vehicles and occlusions. A related system, ROADF (Zlotnick and Carnine, 1993), attempts to automate the seeding process. This system searches for antiparallel sets of edge chains to act as initial road fragments. Higher level linking is then undertaken to join intermittent road sections and to eliminate 'false' seeds.

Neither of the above systems is currently robust enough for 'general purpose' road network retrieval. As with low resolution road finders they work best in simpler landscapes such as rural or low density suburban settings. In denser built-up areas, conflicting cues arise, for example, 'false' seeds associated with parallel building roof and shadow boundaries. Since building detection and road detection methods rely on the similar pre-processing methods, it seems unlikely that extraction systems for either class in isolation will be successful and therefore a comprehensive knowledge base of potentially conflicting objects must be included. The tools for higher level reasoning needed to satisfactorily link disjoint road sections other than by simple mathematical extrapolation do not currently exist although research is underway in this area (e.g. Barzohar and Cooper, 1996).

### 3DIUS (3-Dimensional Image Understanding System)

The previous examples described efforts toward the goal of fully automated object detection and classification. There are, however, parallel developments designed to support conventional human interpretation by addressing 2 key areas, namely, (a) methods for generating fully integrated site models consisting of multiple images and associated collateral data and (b) sophisticated visualization to improve human interpretation (Gee and Newman, 1993). Pioneering work in these areas has been conducted at SRI with the development of the Cartographic Modelling Environment (CME) (Hanson and Quam, 1988). The CME provides tools for spatial integration of diverse data sets such as stereo imagery, DEMs and 3-D object space models and for viewing these data from arbitrary perspective directions. Subsequently, the CME has evolved to 3DIUS a system which incorporates more sophisticated uses of photogrammetry and image rendering. A spin-off research activity is the generation and manipulation of 'virtual' worlds, i.e. landscape databases which encompass fully integrated object models derived from multisensor inputs with conventional geographic and image data. While the principal applications of these databases have been in the areas of military mission planning and dynamic simulation (Fischler and Bolles, 1994; McKeown *et al*., 1994; McKeown *et al*., 1996), civilian applications, such as urban planning, are also being pursued (e.g. Sinning-Meister *et al*., 1996).

## RESEARCH AND DEVELOPMENT REQUIREMENTS

From reviews of the state-of-the-art of digital mapping systems (e.g. Heipke, 1993; Derenyi,1993; Leberl, 1994), it is clear that while robust automated methods for orthorectification and digital elevation model extraction are available, planimetric information extraction remains a primarily operator-intensive operation. On the other hand, aerial IU technology contains methodological ingredients which have the potential to automate at least part of the planimetric feature extraction process. In this section we address a selected set of key image processing issues and propose directions for further research with the objective of developing practical tools for an operational mapping environment. It is recognized that extracting and manipulating planimetric information involves much more than image processing; however such issues as GIS, databases and semantics go beyond the scope of this paper and will not be addressed here.

### Intelligent Image Segmentation

While numerous segmentation strategies and variants have been proposed, most have been applied to relatively limited image datasets, usually of unspecified spatial resolution. A systematic evaluation of the most promising methods is needed, directed specifically at metre-scale resolution imagery and encompassing a broad range of landscapes. Relative and absolute performance measures must be linked to operational criteria such as map production throughput and cartographic accuracy. As an example, the ability of a method to generate segments which exhibit a one-to-one mapping with physical objects or object components is highly desireable since both over-fragmentation or over-generalization pose problems for efficient labelling. Second, while most segmentation algorithms rely solely on statistical decision making, the development of intelligent goal-directed process control, i.e. the tailoring of segmentation parameters to expected scene thematic content, is critical and holds more promise for improved performance than the search for new segmentation and edge detection methods (Matsuyama, 1987).

### Diversification of Functional Area Models

Systems, such as SPAM, which attempt to derive full scene interpretations, employ image reasoning on a hierarchy of structural scales. Critical ingredients in this process are functional area models and knowledge which relate neighbouring objects since only this type of information can resolve conflicting feature classifications which arise when only in-situ characteristics are considered. While some success has been achieved in developing functional spatial models for a restricted class of man-made settings (i.e. airports and simple suburban residential areas), most IU research is directed at the extraction of isolated feature types (e.g. roads, buildings) and therefore do not attempt to exploit thematic context. To be of operational value, IU technology will must include a suite of robust models that encapsulate the structural content of a broad range of 'cultural' settings (e.g. rural-agricultural, urban-industrial, urban-commercial, urban-dense residential) as well as 'natural' landscapes where hydrological-land cover relations are encapsulated. Limited work in natural settings has been already been undertaken, for example with the development of expert systems, such as DNESYS to rationalize drainage networks derived from digital terrain models (Qian *et al*., 1990; Smith *et al*., 1990; Hadipriono *et al*., 1990).

### Exploitation of Collateral Data

Map updating is a key cartographic activity where an 'old' map (here we assume it is digital) forms a source of collateral data. Below are some areas which have automation potential.

(a) Methods are needed to compare and relate abstract map and fragmented image features, the latter derived from an intial image segmentation. Relational graph-based methods are currently under study and show promise for delineating broad areas of apparent 'no-change' and 'significant change' (Stilla, 1995; Moissinac *et al*., 1995).
(b) In areas of 'no-change', the derivation of 'training' attributes for object classes and object functional relationships in order to supplement generic functional models with 'local' knowledge.
(c) In areas of apparent change, procedures to derive revised area classifications, abstract these and integrate them with the existing collateral information to generate revised maps and site models.

### Hierarchical Processing

In addition to high monochrome high spatial resolution imagery, the mapping community will have access to lower resolution multispectral coverage either from other satellites (e.g. SPOT and Landsat) or from an alternate imaging modes. Methods for exploiting moderate resolution (20 to 100 metres) multispectral data to define broad land cover classes are well established, however, further work is needed to refine and extend these to resolutions on the scale of 4 metres. The performance of IU systems could be improved by employing this technology within the context of a spatial processing hierarchy. For example, a single classified multispectral image could provide initial coarse site model while corresponding multitemporal data could provide cues to the location of areas of thematic change. The high resolution imagery would then spatially refine these interpretations. The incorporation and employment of ancillary spectral knowledge in IU systems has to date been limited (Ford and McKeown, 1992) but warrants further consideration. Extensive research has been conducted in pyramidal image processing (i.e. coarse to fine scale structure analysis within a given image), however, to be effectively exploited in IU, these must be integrated with parallel levels of knowledge representations (Silberberg, 1988).

### Perceptual Organization

Perceptual organization is defined as the ability of a visual system to capture visual representations of structure and similarity among otherwise random elements, features and patterns (Price and Huertas, 1992). Pioneering work has been conducted in developing and applying grouping techniques to building detection in single images (Huertas and Nevatia, 1988) and in building reconstruction from stereo images (Dang *et al*., 1994) However, the application potential of this concept is much broader. In the mapping area, models of the spatial organization of urban street networks could be employed to connect fragmentary elements found by conventional road finding algorithms. A second application area is image-map matching. In scenarios where high resolution imagery is to be used to update existing low quality maps, precise image-map registration may be impossible because of poor map geometric accuracy. Rather than using ground control points for registration, control 'groups' could be extracted and matched to provide high precision local registration. A control group would consist of a set of logically associated objects such as road intersections (Haala and Vosselman, 1992; Hu and Pavlidis, 1996).

## CONCLUSIONS

A perceived benefit of operational spaceborne high resolution (1 - 3 metres) imagery will be its utility for detailed mapping worldwide. A processing bottleneck in map generation is currently in the area of planimetric information extraction, a process which remains labour intensive. Information extraction techniques which have been developed for low resolution spaceborne imagery are predominately data driven and are of limited use for this new imagery. On the other hand, extensive research has been conducted in aerial image understanding, primarily for the identification and assessment of a limited set of man-made objects. These techniques make extensive use of model-driven spatial reasoning, an approach which is readily generalized to a broader range of planimetric features. We have presented a review of IU technology and identified a number of areas requiring further development. These include model-driven image partitioning, diversified functional area models, integration with collateral image and textual information and perceptual organization.

Finally, migration to full automation in planimetric information extraction is unlikely to be swift but rather will be a phased process involving the evolution of semi-automated systems with ever increasing useage of computerized processing components. The acceptance of automation by the mapping community will be contingent on the development and application of rigorous performance measures which clearly demonstrate automation as a robust, accurate and cost-effective alternative to operator-intensive processing. Unfortunately, until recently (see Hsieh, 1995) little effort has been expended on developing measures to assess the impact of partial automation on overall system performance.

## REFERENCES

Ballard, D.H. and C.M. Brown, 1982, *Computer Vision*, Prentice-Hall Publishing Company

Barnard, S.T., 1990, "Recent Progress in CYCLOPS: A System for Stereo Cartography", *1990 DARPA Image Understanding Workshop*, pp. 449-455.

Barzohar, M. and D.B. Cooper, 1996, "Automatic Finding of Main Roads in Aerial Images by Using Geometric-Stochastic Models and Estimation", *IEEE Transactions on Pattern Recognition and Machine Intelligence*, Vol. 18, pp. 707-721.

Bhanu, B. and J.C. Ming, 1988, “TRIPLE: A Multi-Strategy Matching Learning Approach to Target Recognition”, *1988 DARPA Image Understanding Workshop*, pp. 537-547.

Bjorklund, C., Noga, M., Barrett, E. and D. Kuan, 1989, “Lockheed Imaging Technology Research for Missiles and Space”, *1989 DARPA Image Understanding Workshop*, pp. 332-352.

Bogdanowicz, J.F. and A. Newman, 1989, “Overview of the SCOPIUS Program”, *1989 DARPA Image Understanding Workshop*, pp. 298-308.

Bremner, B., Hoogs, A. and J. Mundy, 1996, “Integration of Image Understanding Exploitation Algorithms in the RADIUS Testbed”, *1996 ARPA Image Understanding Workshop*, pp. 255-268.

Brooks, R.A., 1981, “Symbolic Reasoning Among 3-D Models and 2-D Images”, *Artificial Intelligence*, Vol. 17, pp. 285-349.

Canny, J., 1986, “A Computational Approach to Edge Detection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, pp. 679-698.

Corby N.R., Mundy, J.L., Vrobel, P.A., Hanson, A.J., Quam, L.H., Smith, G.B. and T.M. Strat, 1988, “PACE An Environment for Intelligence Analysis”, *1988 DARPA Image Understanding Workshop*, pp. 342-350.

Dang, T., Jamet, O. and H. Maitre, 1994, “Applying Perceptual Grouping and Surface Models to the Detection and Stereo Reconstruction of Buildings in Aerial Images”, *Proceedings of the 1994 ISPRS Commission III Symposium*, pp. 165-172.

Derenyi, E.E., 1993, “Low Cost Soft Copy Mapping”, *Procceedings of the SPIE*, Vol. 1944, pp. 223-230.

Dissard, O. and O. Jamet, 1995, “3-D Reconstruction of Buildings from Stereo Images Using Both Monocular Analysis and Stereo Matching: An Assessment Within the Context of Cartographic Production”, *Proceedings of the SPIE*, Vol. 2486, pp. 255-266.

Edwards, J., Gee, S., Newman, A., Onishi, R., Parks, A., Sleeth, M. and F. Vilnotter, 1992, “RADIUS: Research and Development for Image Understanding Systems Phase 3”, *1992 DARPA Image Understanding Workshop*, pp. 177-176.

Fischler, M.A. and R.C. Bolles, 1994, “Image Understanding Research at SRI International”, *1994 DARPA Image Understanding Workshop*, pp. 53-68.

Ford, S.J. and D.M. McKeown, 1992, “Utilization of Multispectral Imagery for Cartographic Feature Extraction”, 1*992 DARPA Image Understanding Workshop*, pp. 805-820.

Forstner, W., 1993, “Feature Extraction in Digital Photogrammetry”, *Photogrammetric Record*, Vol. 14, pp. 585-611.

Fua, P. and A.J. Hanson, 1988, “Extracting Generic Shapes Using Model-Driven Optimization”, *1988 DARPA Image Understanding Workshop*, pp. 994-1004.

Fung, K.B., Goodenough, D.G., Forde, B., Robson, M.A., Kushigbor, C., Guo, A., Menard, A., Dostie, M. and T. Fowlow, 1993, “The System of Hierarchical Experts for Resource Inventories (SHERI)” *Proceedings of the 16th Canadian Symposium on Remote Sensing*, pp. 879-884.

Gee, S.J. and A.M. Newman, 1993, “Conceptual Design for a Model-Supported Exploitation Workstation for Imagery Analysis”, *Proceedings of the SPIE*, Vol. 1944, pp. 242-253.

Goodenough, D.G., Goldberg, M., Plunkett, G. and J. Zelek, 1987, “An Expert System for Remote Sensing”, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 25, pp. 349-359.

Haala, N. and G. Vosselman, 1992, “Recgonition of Road and River Patterns by Relational Matching”, *ISPRS Commission III Archives*, Vol. 29, pp. 969-975.

Hadipriono, F.C., Lyon, J.G., Li, T. and D.P. Argialas, 1990, “The Development of a Knowledge-Based Expert System for Analysis of Drainage Patterns”, *Photogrammetric Engineering and Remote Sensing*, Vol. 56, pp. 905-909.

Hanson, A.J. and L.H. Quam, 1988, “Overview of the SRI Cartographic Modeling Environment”, *1988 DARPA Image Understanding Workshop*, pp. 576-582.

Haralick, R.M. and L.G. Shapiro, 1992, *Computer and Robot Vision*, Addison-Wesley Publishing Company.

Haralick, R.M., 1979, “Statisical and Structural Approaches to Texture”, *Proceedings of the IEEE*, Vol. 67, PP. 786-804.

Harvey, W.A., Diamond, M. and D.M. McKeown, 1992, “Tools for Acquiring Spatial and Functional Knowledge in Aerial Image Analysis”, *1992 DARPA Image Understanding Workshop*, pp. 857-873.

Harvey, W.A. and M. Tambe, 1993, “Experiments in Knowledge Refinement for a Large Rule-Based System”, *Technical Report CMU-CS-93-195*, Carnegie Mellon University, 14 pages.

Heipke, C., Englisch, A., Speer, T., Steir, S. and R. Kutka, 1994, “Semi-Automatic Extraction of Roads from Aerial Images”, *ISPRS Commission III Archives*, Vol. 30, pp. 353-360.

Heipke, C., 1993, “State-of-the-Art of Digital Photogrammetric Workstations for Topgraphic Applications”, *Proceedings of the SPIE*, Vol. 1944, pp. 210-222.

Hsieh, Y.C., McKeown, D.M. and F.P. Perlant, 1992, “Performance Evaluation of Scene Registration and Stereo Matching for Cartographic Feature Extraction”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, pp. 214-237.

Hsieh, Y., 1995, “Design and Evaluation of a Semi-Automated Site Modelling System”, *Technical Report CMU-CS-95-195*, Carnegie Mellon University, 76 pages.

Hu, J. and T. Pavlidis, 1996, “A Hierarchical Approach to Efficient Curvilinear Object Searching”, *Computer Vision and Image Understanding*, Vol. 63, pp. 208-220.

Huertas, A. and R. Nevatia, 1988, “Detecting Buildings in Aerial Images”, *Computer Vision, Graphics and Image Processing*, Vol. 41, pp. 131-152.

Irwin, R.B. and D.M. McKeown, 1989, “Methods for Exploiting the Relationship Between Buildings and Their Shadows in Aerial Imagery”, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 19, pp. 1564-1575.

Konecny, G., 1996, “Current Status and Future Possibilities for Topographic Mapping from Space”, *Proceedings of the 14th EARSel Symposium*, pp. 1-18.

Leberl, F.W., 1994, “Practical Issues in Softcopy Photogrammetric Systems”, *ASPRS Conference on Mapping and Remote Sensing Tools for the 21st Century*, pp. 223-230.

Leclerc, Y.G., 1988, “Constructing Simple Stable Descriptions for Image Partitioning”, *1988 DARPA Image Understanding Workshop*, pp. 365-382.

Leclerc, Y.G., 1989, “Image and Boundary Segmentation via Minimal-Length Encoding on the Connection Machine”, *1989 DARPA Image Understanding Workshop*, pp. 1056-1069.

Lin, C. and R. Nevatia, 1996, “Buildings Detection and Description from Monocular Aerial Images”, *1996 DARPA Image Understanding Workshop*, pp. 461-468.

Liow, Y.-T. and T. Pavidis, 1990, “Use of Shadows for Extracting Buildings in Aerial Images”, *Computer Vision, Graphics and Image Processing*, Vol. 49, pp. 242-277.

Marr, D. and H. Hildreth, 1980, “Theory of Edge Detection”, *Proceedings of the Royal Society* London Series B, Vol. 207, PP.187-217.

Matsuyama, T., 1987, “Knowledge-Based Aerial Image Understanding

Systems and Expert Systems for Image Processing", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 25, pp. 305-316.

McConnell, C.C. and D.T. Lawton, 1988, "IU Software Environments", *1988 DARPA Image Understanding Workshop*, pp. 666-677.

McGlone, J.C. and J.A. Shufelt, 1993, "Incorporating Vanishing Point Geometry in Building Extraction Techniques", *Proceedings of the SPIE*, Vol. 1944, pp. 273-284.

McGlone, J.C. and J.A. Shufelt, 1994, "Projective and Object Space Geometry for Monocular Building Extraction", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.54-61.

McKeown, D.M., Harvey, W.A. and J. McDermott, 1985, "Rule-Based Interpretation of Aerial Imagery", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 7, pp. 570-585.

McKeown, D.M. and J.L. Denlinger, 1988, "Cooperative Methods For Road Tracking in Aerial Imagery", *1988 DARPA Image Understanding Workshop*, pp. 327-341.

McKeown, D.M., Harvey, W.A. and L.E. Wixson, 1989, "Automating Knowledge Acquisition for Aerial Image Interpretation", *Computer Vision, Graphics and Image Processing*, Vol. 46, pp. 37-81.

McKeown, D.M. and F.P. Perlant, 1992, "Refinement of Disparity Estimates Through the Fusion of Monocular Image Segmentations", *1992 DARPA Image Understanding Workshop*, pp. 839-855.

McKeown, D.M., *et al*., 1994, "Research in the Automated Analysis of Remotely Sensed Imagery: 1993-1994", *1994 DARPA Image Understanding Workshop*, pp. 99-132.

McKeown, D.M., Gifford, S.J., Polis, M.F., McMahill, J. and C.D. Hoffman, 1996, "Progress in Automated Virtual World Construction", *Technical Report CMU-CS-96-102*, 13 pages.

Medioni, G. and Y. Yasumoto, 1987, "Corner Detection and Curve Representation Using Cubic B-Splines", *Computer Vision, Graphics and Image Processing*, Vol. 39, pp. 267-278.

Mohan, R. and R. Nevatia, 1988, "Perceptual Grouping for the Detection and Description of Structures in Aerial Images", *1988 DARPA Image Understanding Workshop*, pp. 512-526.

Mohan, R. and R. Nevatia, 1989, "Using Perceptual Organization to Extract 3-D Structure"', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, pp. 1121-1139.

Moissinac, H., Maitre, H. and I. Bloch, 1995, "Graph Based Urban Scene Analysis Using Symbolic Data", *Proceedings of the SPIE*, Vol. 2486, pp. 93-104.

Mueller, W.J. and J.A. Olson, 1993, "Model-Based Feature Extraction", *Proceedings of the SPIE*, Vol. 1944, pp. 263-272.

Mundy, J.L. *et al*., 1992, "The Image Understanding Environment Program", *1992 DARPA Image Understanding Workshop*, pp. 185-214.

Murphy, M.E., 1993, "Rapid Generation and Use of 3D Site Models to Aid Imagery Analysts/Systems Performing Image Exploitation", *Proceedings of the SPIE*, Vol. 1944, pp. 254-262.

Nevatia, R. and K.R. Babu, 1980, "Linear Feature Extraction and Description", *Computer Vision, Graphics and Image Processing*, Vol. 13, pp. 257-269.

Nicolin, B. and R. Gabler, 1987, "A Knowledge-Based System for the Analysis of Aerial Images", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 25, pp. 317-329.

Pal, N.R. and S.K. Pal, 1993, "A Review of Image Segmentation Techniques", *Pattern Recognition*, Vol. 26, pp. 1277-1294.

Pan, H.-P., 1994, "Two-Level Global Optimization for Image Segmentation", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 49, pp. 21-32.

Perkins, W.A., Laffey, T.J. and T.A. Nguyen, 1985, "Rule-base Interpreting of Aerial Photographs using LES", Proceedings of the SPIE, Vol. 548, pp. 138-146.

Perkins, W.A., 1980, "Area Segmentation of Images Using Edge Points", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 2, pp. 8-15.

Price, K. and A. Huertas, 1992, "Using Perceptual Grouping to Detect Objectsin Aerial Images", *ISPRS Commission III Archives*, Vol. 29, pp. 842-855.

Qian, J., Ehrich, R.W. and J.B. Campbell, 1990, "DNESYS - An Expert System for Automatic Extraction of Drainage Networks from Digital Elevation Data", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, pp. 29-45.

Reed, T.R. and J.M. H. du Buf, 1993, "A Review of Recent Texture Segmentation and Feature Extraction Techniques", *Computer Vision, Graphics and Image Processing: Image Understanding*, Vol. 57, pp. 359-372.

Roux, M. and D.M. McKeown, 1994, "Feature Matching for Building Extraction from Multiple Views", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 46-53.

Roux, M., Hsieh, Y.C. and D.M. McKeown, 1995, "Performance Evaluation of Object Space Matching for Building Extraction Using Several Images", *Proceedings fo the SPIE*, Vol. 2486, pp.277-297.

Sarjakoski, T., 1988, "Artificial Intelligence in Photogrammetry", *Photogrammetria*, Vol. 42, pp. 245-270.

Shi, Z. and R. Shibasaki, 1996, "Towards Automated House Detection from Digital Stereo Imagery for GIS Databases", *ISPRS Commission IV Archives*, Vol. 31, pp. 780-785.

Shufelt, J.A. and D.M. McKeown, 1993, "Fusion of Monocular Cues to Detect Man-Made Structures in Aerial Imagery", *Computer Vision, Graphics and Image Processing*, Vol. 57, pp. 307-330.

Sinning-Meister, M., Gruen, A. and H. Dan, 1996, "3D City Models for CAAD-Supported Analysis and Design of Urban Areas", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 51, pp. 196-208.

Silberberg, T.M., 1988, "Multiresolution Aerial Image Interpretation", *1988 DARPA Image Understanding Workshop*, pp. 505-511.

Smith, T.R., Zhan, C. and P. Gao, 1990, "A Knowledge-Based, Two-Step Procedure for Extracting Channel Networks from Noisy DEM Data", *Computers and Geosciences*, Vol. 16, pp. 777-786.

Sonka, M., Hlavac, V. and R. Doyle, 1993, *Image Processing, Analysis and Machine Vision*, Chapman and Hall Computing Series.

Stilla, U., 1995, "Map-Aided Structural Analysis of Aerial Images", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 50, pp. 3-10.

Venkateswar, V. and R. Chellappa, 1992, "Hierarchical Stereo Matching using Feature Groupings", *1992 DARPA Image Understanding Workshop*, pp. 427-437.

Wang, J. and W. Liu, 1994, "Road Detection from Multi-Spectral Satellite Imagery", *Canadian Journal of Remote Sensing*, Vol. 20, pp. 180-191.

Xu, Y., 1992, "A Conception for Object-Oriented Description of Knowledge and Data in Scene Analysis and Object Recognition", *ISPRS Commission III Archives*, Vol. 29, pp.220-226.

Zelek, J.S., 1990, "Computer-Aided Linear Planimetric Feature Extraction", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, pp. 567-572.

Zlotnick, A. and P.D. Carnine, 1993, "Finding Road Seeds in Aerial Images", *Computer Vision, Graphics and Image Processing: Image Understanding*, Vol. 57, pp. 243-260.