GEOLOGICAL SURVEY OF CANADA
OPEN FILE 6752

Automatic Image Analysis of Sediment Types: Mapping from Georeferenced Video Footage on the Labrador Shelf

K. Jerosch, A. Lüdtke, P. Pledge, O. Paitich, V.E. Kostylev

2011

Natural Resources Canada   Ressources naturelles Canada

GEOLOGICAL SURVEY OF CANADA
OPEN FILE 6752

Automatic Image Analysis of Sediment Types: Mapping from Georeferenced Video Footage on the Labrador Shelf

K. Jerosch[1], A. Lüdtke[2], P. Pledge[1], O. Paitich[1], V.E. Kostylev[1]

[1] Geological Survey of Canada (Atlantic), Bedford Institute of Oceanography, P.O. Box 1006, Dartmouth, Nova Scotia B2Y 4A2

[2] Center for Computing and Communication Technologies (TZI), University Bremen, Am Fallturm 1, 28359 Bremen, Germany

2011

doi:10.4095/ 288055

This publication is available from the Geological Survey of Canada Bookstore (http://gsc.nrcan.gc.ca/bookstore_e.php).
It can also be downloaded free of charge from GeoPub (http://geopub.nrcan.gc.ca/).

Recommended citation:
Jerosch, K., Lüdtke, A., Pledge, P., Paitich, O. and Kostylev, V.E., 2011. Automatic Image Analysis of Sediment Types: Mapping from Georeferenced Video Footage on the Labrador Shelf; Geological Survey of Canada, Open File 6752, 36 p. doi:10.4095/ 288055

Publications in this series have not been edited; they are released as submitted by the author.

Table of Contents

Lists of Figures II

Lists of Tables III

1 Introduction 3

2 Data Acquisition - Video Footage Labrador Shelf 2010 4

3 Methods 6

3.1 Preparing the Data Sets 6

3.1.1 Trimming and Metadata of Video Footage 6

3.1.2 Scale 6

3.1.3 Navigation 7

3.2 Georeferencing Still Photographs 7

3.3 Extracting Georeferenced Frames From Video Footage 9

3.4 *Gidas* - Geospatial Image Database and Analysis System 10

3.4.1 Getting Started 11

3.4.2 Data Import and Preparation 12

3.4.3 Annotation and Feature Export of Training Data Set 13

3.4.4 Machine Learning Toolkit *Weka* 14

3.4.5 Model Application on *Gidas*: Annotations and Predictions 15

3.4.6 GIS Performance of the Results 15

3.5 *Gidas* - Guideline 16

3.5.1 Getting Started 16

3.5.2 Data Import and Preparation 17

3.5.3 Annotation and Feature Export of Training Data Set 17

3.5.4 *Weka* 18

3.5.5 Model Application on *Gidas*: Annotations and Predictions 19

3.5.6 *Gidas* Database Backup 20

3.5.7 GIS Performance of the Results 20

4 Results for the Labrador Shelf 21

4.1 Georeferenced Images Extracted From HD Video Footage Displayed in 3D 21

4.2 Automated Detection of Sediment Types on Georeferenced Imagery 23

4.2.1 Methodological Settings: Classification Scheme and Grid Sizes 23

4.2.2 *Weka* - Summary 24

4.2.3 *Weka* - Detailed Accuracy By Class 25

4.2.4 *Weka* - Confusion Matrix 25

4.2.5 GIS Performance 26

5 Summary and Discussion 30

6 Acknowledgements 32

7 References 33

List of Figures

Figure 1
The red rectangle highlights the area of investigation on the Labrador Shelf. 4

Figure 2
GSCA DeepImager system. 4

Figure 3
Typical seabed still images of sandy and gravelly areas from the Labrador Shelf which were extracted from HD video footage. The resolution of these images allows identification of benthic species and sediment types. 5

Figure 4
The *JPEG Merge Tool*. 7

Figure 5
The *Image-Georeferencing-Tool*. 8

Figure 6
Example how to apply the *Video-Frame-Georeferencing-Tool*. 9

Figure 7
Gidas screen shot. A generic grid is the basis for the annotation of sea floor features. 11

Figure 8
Weka screen shot showing the feature distribution. 14

Figure 9
The video transect visualized as a 285 m red line in distance on the 3D multibeam. It crosses a 3–5 m deep and 20–30 m wide fresh ice scour on the Labrador Shelf (see Figure 6). 21

Figure 10
Zooming slightly into (A) so-called georeferenced images can be seen on the transect (navigation shown as red dots). Zooming in even more (B) enables the recognition of features like rocks in the images. 22

Figure 11
Application of the Wentworth (1922) sediment classification system in *Gidas* (above). Scale is given by the overlay of a 15 x 15 cm grid in UTM 21N (WGS84) coordinates. The lower image shows the non-annotated image. 23

Figure 12
GIS performance of the results. Upper three panels: Annotations on training data sets: manually determined image areas being either “mud or sand” (orange), “granules or pebbles” (dark blue), “cobbles” (red), “boulders” (yellow), and “urchins” (green). Black color determines negligible (“don‘t care“) and light blue unannotated images areas. Lower three panels: Results of the prediction model developed by *Weka* applied to the remaining data set by means of *Gidas*. 27

Figure 13
Detailed GIS performance of the results.Upper three panels: Annotations on training data sets. The manually determined image areas being either “mud or sand” (orange), “granules or pebbles” (blue), “cobbles” (red), “boulders” (yellow), and “urchins” (green). The class “don‘t care” in black determines negligible image areas. Lower three panels: Prediction model developed by *Weka* applied to the remaining data set by means of *Gidas* overlain onto georeferenced video frames. 28

Figure 14
3D-visualization of the 15 cm grid prediction results. 29

Figure 15
Summary of the method. 31

List of Tables

Table 1
Statistical mean values resulting from the application of a machine learning model to the three (7-, 15- and 25-cm-grid) training data sets given in the summary of the *Weka* software. The general trend of the prediction quality is given by absolute numbers and percentages of correct and incorrect classified cells after a cross-validation. ... 24

Table 2
Metrics to thevaluate the correctness of a pattern recognition algorithm: Precision can be seen as a measure of exactness or fidelity, whereas recall is a measure of completeness. True positive (TP) is the number of items correctly labeled as belonging to the positive class; false positives (FP) are items incorrectly labeled as belonging to the class. F-measure is a weighted harmonic mean of precision and recall and most suitable for quality estimation and comparison. 25

Table 3
Confusion Matrix (a: mud or sand and brittle stars; b: granules or pebbles; c: cobbles; d: boulders; e: urchins; f: don't care). For a comparison the class sizes are given for each class and grid cell sizes showing the distribution of the instances ("Total of Instances" and "% of Total"). The best distribution would be evenly distributed instances in order to avoid weighted results. Percentages of incorrect predicted instances by class are given in the column "% Wrong"........ 26

1 Introduction

One of the major objectives in the field of automatic image analysis when applied to marine sciences is to detect, identify, and track sea floor characteristics in underwater video footage acquired by fixed or towed cameras. This includes the possibility of using these techniques for future work on Remotely Operated Vehicles (ROVs) and Autonomous Underwater Vehicles (AUVs). Recently, the analysis of still photographs and video clips in marine sciences has mostly been performed manually by investigation of a determined subset of the data available (e.g., Kostylev et al., 2001). Automatic feature detection has been applied in marine sciences but the implementation of suitable algorithms is particularly complicated (e.g. Guinan et al., 2009, Jerosch et al., 2006, Lüdtke et al., accepted, Purser et al., 2009). If video data comes from a towed system, algorithms have to deal with inadequate battery-powered lighting, and varying speed, pitch, roll and altitude of the camera above the sea floor. The quality of a video stream coming from a ship-powered ROV is enhanced by a more stable speed, more constant height over ground, and more homogenous lighting conditions. What both systems have in common is that the algorithms have to track a broad spectrum of known and unknown features living on the sea floor.

Digital image processing provides powerful tools for fast and precise analysis of large image data sets in marine and geoscientific applications. Facing the increasing amount of georeferenced image and video data acquired by ROVs, AUVs, and towed systems, a method of automatic image analysis is required. A new and rapidly-evolving application is the combination of video footage and georeferenced frame extraction for sea floor habitat mapping.

This Open File describes mapping sea floor features efficiently from video footage providing a guideline for every required step. This is achieved by

1. developing a method to extract single frames from the video footage and make them suitable for Geographical Information Systems (GIS) (georeferencing the frames automatically), and
2. applying the Geospatial Image Database and Analysis System *Gidas* (Lüdtke et al., accepted) on HD video frames from the Labrador Shelf, *Gidas* provides an approach for a fully automatic detection and the quantification of sediment texture on sea floor images/frames on the Labrador Shelf. This is performed based on the extraction of textural features from the images and a statistical classification of these features using machine learning techniques. The approach includes a quality control based on a crossvalidation of the predictions.
3. Furthermore, the mapping of the results and discussing *Gidas* of being a suitable technique for sedimentological seabed mapping forms the concluding scope of the Open File report.

3 Data Acquisition - Video Footage Labrador Shelf 2010

The Labrador Shelf seabed is known for numerous curvilinear iceberg scours and sub circular iceberg-generated pits (Barrie, 1980, Bass and Woodworth-Lynas, 1988, Todd et al., 1988). The scouring potential of an iceberg depends on several factors like its size, shape, drift velocity, and the drag coefficient. The scour size depends on the sediment type, its shear strength, and the bathymetry (Charia et al., 1980).

To study the effect of icebergs on the seabed, six video transects were recorded on the Labrador Shelf (CCGS *Matthew* 2010–037) with the following system and settings:

Camera Name:	GSCA DeepImager
Video Resolution:	HD 1080i
Still Image Resolution:	12.0 MegaPixel
Pressure case viewport:	Optically corrected dome port
System Pressure rating:	4000 m tested. MBES 300 m (Multi beam Echo Sounder rated for a pressure of 300 m)
Scale Lasers:	2 x 532 nm 5 mW green lasers
Attitude Sensor:	HMR3000 (heading, pitch and roll)
Pressure Sensor:	Valeport IPS
MultiBeam:	Imagenex 837 DeltaT (120 degree swath)
Max. Dive Time:	2 hours on a single charge (tested)
Typical Operation:	Vessel drifts with the system suspended 1–3 m from the seabed

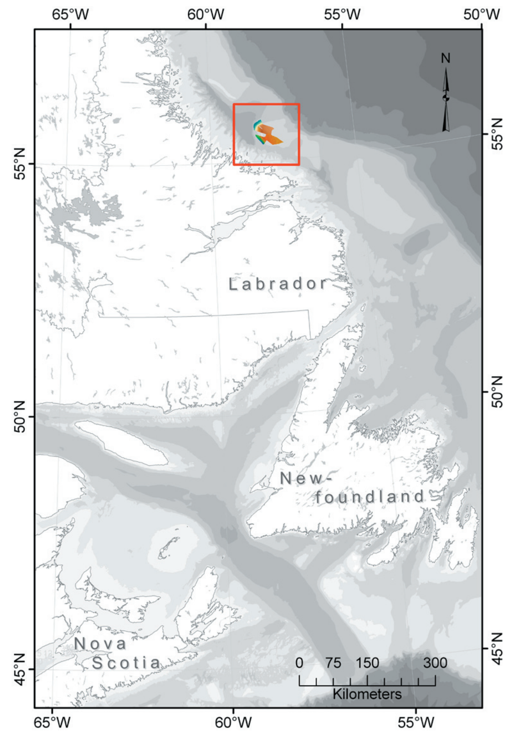


Figure 1 The red rectangle highlights the area of investigation on the Labrador Shelf.

The GSCA DeepImager camera system (Figure 2) is mounted on an integrated sled which successfully completed shallow water testing (up to 280 m) on the CCGS *Matthew* on two cruises (Makkovik Bank, St. Anne's Bank). It is deployed from the stern of the vessel and lowered to the seabed. As the camera system approaches the seabed, individual components are turned on (lights, cameras, etc.) in order to minimize power consumption and maximize bottom time.

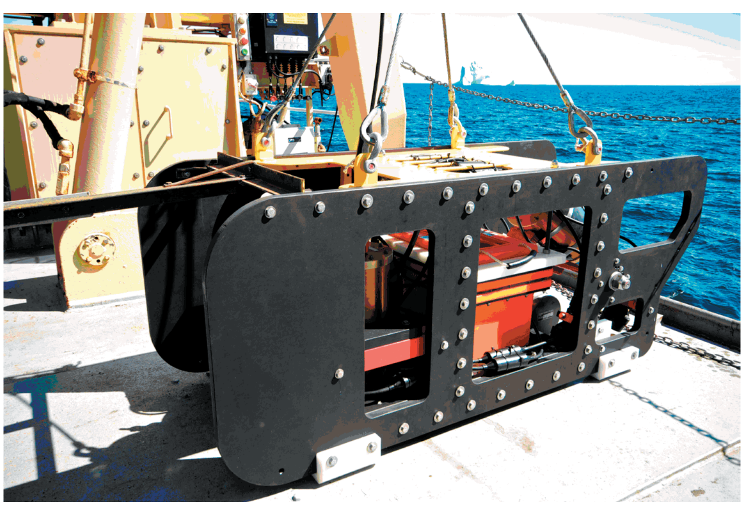


Figure 2 GSCA DeepImager system.

The HD camera is a downward-looking vertical camera which shoots through an optically-corrected domed viewport to eliminate distortion and magnification associated with other flat port systems. A viewfinder image of what the HD camera sees is viewable on the vessel in real time. This view is used to control the camera system's altitude above the seabed by means of an operator-controlled remote winch, to judge focus, lighting, etc.

The HD camera is also capable of shooting still images while recording HD video. The operator can shoot individual stills or set the system to automatically collect stills every 15 seconds (recommended mode). Both stills and HD video are downloaded when the camera system returns to the surface. HD video, still imagery and multibeam data is georeferenced upon completion of the dive.

Figure 3 shows typical seabed still images of sandy and gravelly areas from the Labrador Shelf which were extracted from HD video footage. The resolution of these images (see section 3.3) allows identification of benthic species and sediment types. For this Open File we have used video data crossing a fresh ice scour (CCGS *Matthew* 2010–037, station 6).

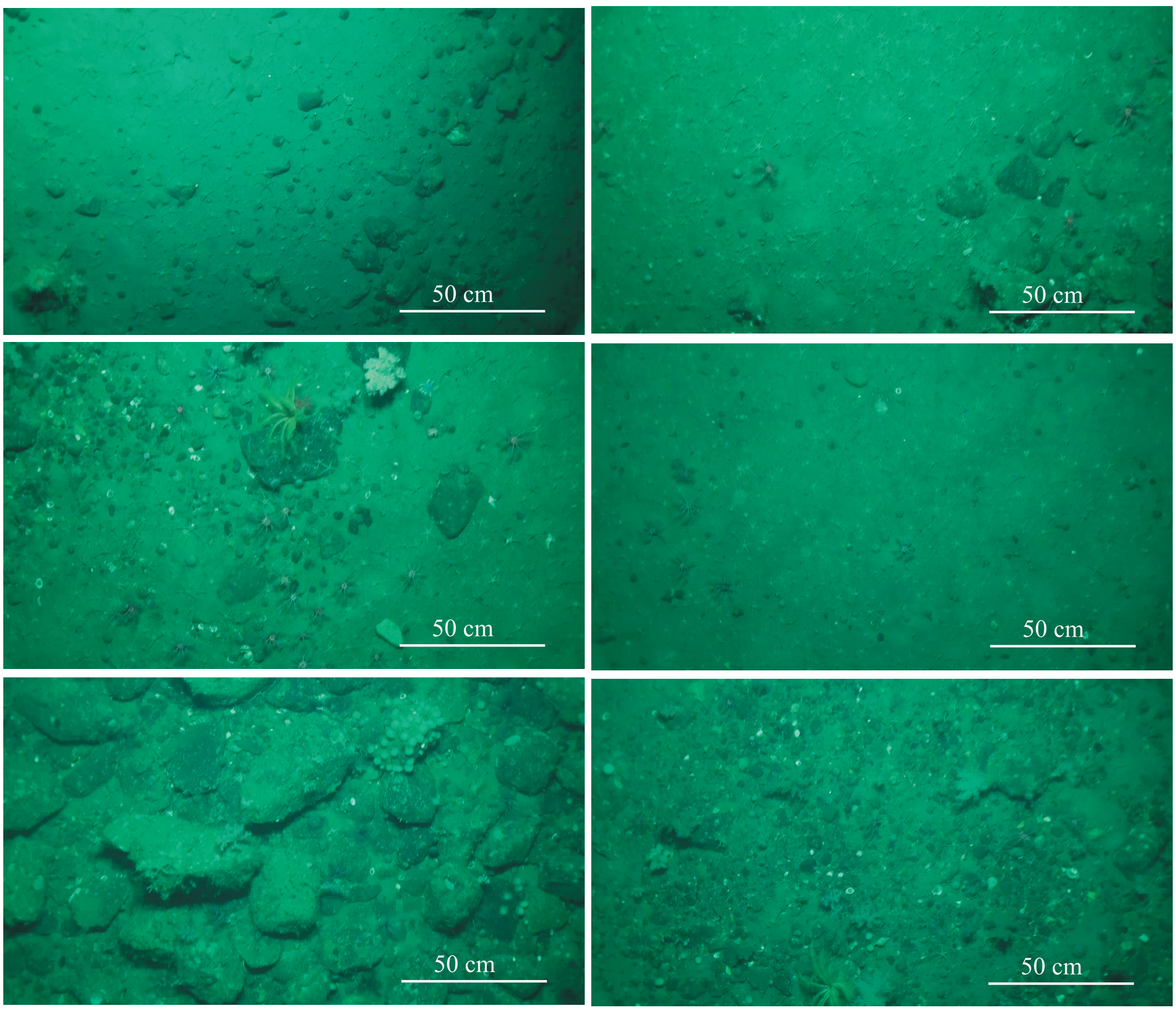


Figure 3 Typical seabed still images of sandy and gravelly areas from the Labrador Shelf which were extracted from HD video footage. The resolution of these images allows identification of benthic species and sediment types.

3 Methods

3.1 Preparing the Data Sets

3.1.1 Trimming and Metadata of Video Footage

The HD videos contain many minutes of data that is unusable (deck footage, falling through the water, coming out of the water, etc.); video with the extension *.m2ts have to be trimmed into usable data (only shots of the ocean floor) using video-editing software such as *Picture Motion Browser*. Depending on the length of the transect, there is usually more than 1 file (*.m2ts) per transect.

Metadata was extracted containing information in 1 second intervals about the transect (date, time stamp, camera settings) as an *.srt file using *DVMP Pro 5*.

*.srt file example

```
1
00:00:00,000 --> 00:00:00,700
          11:32:44    60    12dB  AWB
          05/09/2009  f1.8  0.10  ~ON
00:00:00;00           AUTO  AUTO
2
00:00:00,700 --> 00:00:01,701
          11:32:45    60    12dB  AWB
          05/09/2009  f1.8  0.30  ~ON
00:00:00;21           AUTO  AUTO
3
00:00:01,701 --> 00:00:02,702
          11:32:46    60    12dB  AWB
          05/09/2009  f1.8  0.90  ~ON
00:00:01;21           AUTO  AUTO
```

The information contained within the HD metadata is required externally as an *.srt file for the production of the time stamp and the connection with navigation data. An overlay of the metadata is only displayed in the video when the metadata (*.srt) and video (*.m2ts, *.avi, and others) have the same name and are located in the same folder.

3.1.2 Scale

Georeferenced frame extraction can be used on both HD (*.m2ts) and Campod videos (*.avi) as long as the video start and end times are known. In order to extract them using the *Video-Frame-Georeferncing-Tool*, the pixel size must be defined in metres (assuming the pixel height is identical to the pixel width).

Calculation of the video scale in meters is done using knowledge of the distance between lasers (10 cm):

457 pixels ↔ 10 cm = 0.1 m

1 pixel ↔ (0.1 / 457) m = 0,0012642225031605563 m

Freeware such as *Gimp* or *ImageJ* is used to measure the number of pixels between the camera lasers at the beginning, middle and end of the video. The averaged pixel size is applied to the whole station.

3.1.3 Navigation

Navigation files are made from the files produced by the Track Point. Track Point is a USBL (Ultra Short Base Line) subsea acoustic positioning device. Typically, gear which is either lowered or towed from a vessel will use USBL to position the instrument relative to the vessel. The vessel is equipped with a Trackpoint USBL system and the lowered target would be equipped with a Track Point Beacon. The system sends out an acoustic ping which is received by the beacon on the target. The target beacon responds and this response is detected by the USBL system on the vessel. The response ping is timed and corrected for the speed of sound in water thus giving a range to the target. The USBL system on the vessel also has three transducers in the head aligned very close together (hence the term "Ultra Short Baseline") which are used to triangulate the response ping thus deriving a bearing to the target. Final real-world positioning is further derived by knowing where on the vessel the USBL system is installed.

Track Point documents timestamp, geolocation and sometimes depth. Both the movie and navigation files must have at least the same starting point (time stamp). Navigation must be in the following format: time y x (Note: tab-separated and without header):

- time: OriginalDates hours minutes seconds as long integer: ODHHMMSS (e.g. 248092412)
- y: latitude in UTM coordinates in meters e.g., 4897298.22
- x: longitude in UTM coordinates in meters e.g., 690325.73

Example of a navigation file: OriginalDatesTime, y, x (day-of-year 240, beginning time code of 11:58:50):

240115850	*4911909.452*	*665504.3892*
240115852	*4911905.084*	*665498.677*
240115854	*4911903.72*	*665493.128*
240115900	*4911898.815*	*665492.5344*

3.2 Georeferencing Still Photographs

GSC NavNet jpeg merge V1.2

Still photographs coming from Campod have to be encoded with their spatial location in the image (jpeg) header. This has to be done with *GSC NavNet jpeg merge V1.2* (Figure 4) which embeds the geographical and UTM coordinates from Trackpoint navigation files into the headers of the jpeg exchangeable image file format (Exif). It also creates a *.csv file that is used to load the positioning information into ED (GSC Expedition Database).

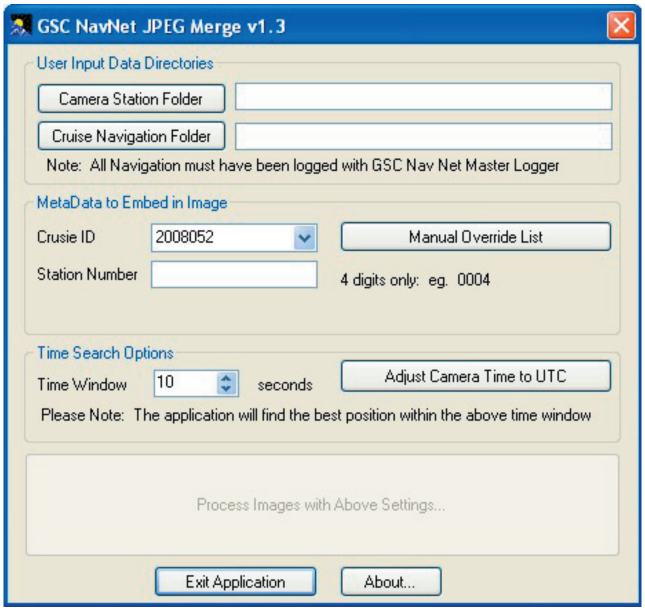


Figure 4 The *JPEG Merge Tool.*

E.g.: *Cruise ID: 2009039 - Stn: 0073 Position: Lat: 44.214952, Lon: -66.610855, E: 690856.29, N: 4898523.31, zone: 19*

IrfanView

In order to use *Gidas,* the high resolution still photographs must be resized with *IrfanView* before georeferencing to avoid long processing times. The following series of operations are performed on each folder containing images:

- File - Batch Conversion/Rename
- Work as Batch Conversion
- Advanced
- Resize - Set new size - set long side to 1024
- Add all images in a folder
- Start Batch

Image-Georeferencing-Tool

Open *GenerateWorldFiles.jar* (Figure 5) and enter three parameters: the path to the folder containing the images, the width, and the height of a pixel in metres. Usually the width and height of a pixel have identical values.

The pixel sizes can be determined as follows:

- same calculation of the image scale in metres as for the videos (see section 3.1.3)
- pixelWidth = pixelHeight = <laser distance in metres> / <laser distance in pixel>

  0.0012642225031605563 m = 0.1 m / 457

Note: In order to determine a Course over Ground (COG) the tool sorts the images alphabetically.

The tool uses the UTM coordinates (e.g., E: 690856.29, N: 4898523.31, zone: 19) coming from the jpg header after applying *GSC NavNet jpeg merge V1.2*. to generate a world file for each image in the folder.

This tool is a result of the cooperation with the TZI (Germany).

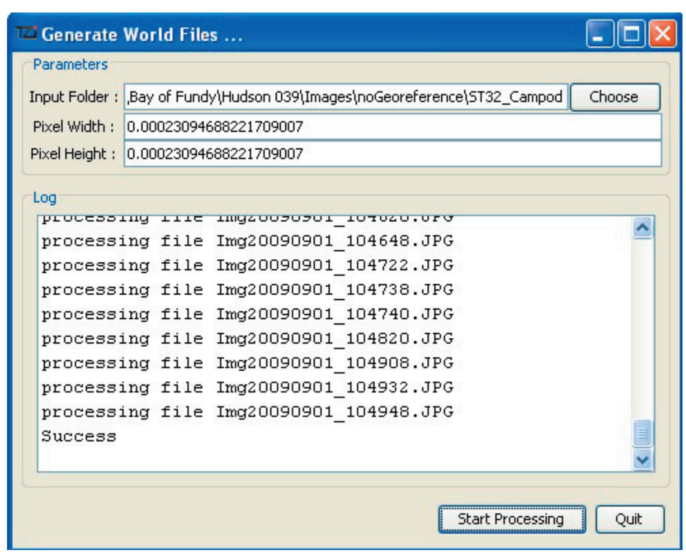


Figure 5 The *Image-Georeferencing-Tool.*

3.3 Extracting Georeferenced Frames From Video Footage

Providing trimmed video footage and navigation as well as metadata (averaged scale and heading) are the last steps to generate georeferenced frames from videos using the *Video-FrameGeoreferencinging-Tool* (Figure 6).

Running the tool - Guideline

1. Include system variables path (location of the tool): e.g., F:\VideoFrameGeoreferencingTool

- ‘My Computer’ right click ‘Properties’
- ‘System Properties’, ‘Advanced’ click ‘Environment Variables’
- choose ‘edit’ the variable ‘path’ in the ‘System Variables’
- add the path to the tool location at the end - Note: avoid space characters after the semicolon

Note: you just have to do this once before the first usage

2. Open panel

- ‘Start’
- ‘Run’
- type in open: ‘cmd’

3. Usage

- type in simple command line interface for the software library

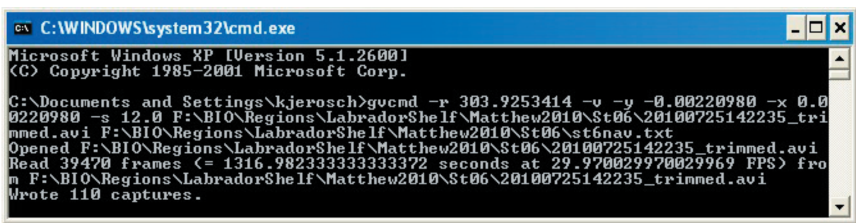


Figure 6 Example how to apply the *Video-Frame-Georeferencing-Tool*.

Handling:

gvcmd [-r <f>] [-T <f>] [-t <f>] [-y <f>] [-x <f>] [-s <f>] [-v] ...
[--] [--version] [-h] <video> <posfile>

Where:

-r <f>, --rotation <f>

camera rotation in radian (Radian = measured angle x 2 x 3.14159265 / 360)

-T <f>, --length <f>
length of the video in seconds

-t <f>, --start <f>
Number of seconds to skip at the beginning of the video file.

-y <f>, --scaley <f>
Pixel size in the y-direction in map units.

-x <f>, --scalex <f>
Pixel size in the x-direction in map units.

-s <f>, --skip <f>
Number of seconds to skip between frames.

-v, --verbose (accepted multiple times)
Increase verbosity

--, --ignore_rest
Ignores the rest of the labeled arguments following this flag.

--version
Displays version information and exits.

-h, --help
Displays usage information and exits.

<video>
(required) Video file from which to read.

<posfile>
(required) Whitespace-separated or tab CSV file from which to read way points.

HELP

gvcmd --help

EXAMPLES

gvcmd -v -y -0.0012642225031605563 -x 0.0012642225031605563 -s 0.1 F:\TZI\VideoFrameGeoreferencing-Tool\Test\ST73_VerticalCam\St73_extr.avi F:\TZI\VideoFrameGeoreferencingTool\Test\ST73_VerticalCam\JD248_UTM.txt

gvcmd -r 5.2883 *-v -y -0.00220980 -x 0.00220980 -s 12.0 F:\BIO\Regions\LabradorShelf\Matthew2010\St06\20100725142235_trimmed.avi F:\BIO\Regions\LabradorShelf\Matthew2010\St06\st6nav.tx*

4. Results

Results consisting of video frame jpegs and their respective world files are stored in the same folder as the video footage file. They can be imported into a GIS, be overlaid on multibeam data, e.g., in *ArcScene,* or used in other image software for further (automatic) analysis such as *Gidas* that requires georeferenced images.

20100725142235_trimmed.avi
20100725142235_trimmed.avi_0.jgw
20100725142235_trimmed.avi_0.jpg
20100725142235_trimmed.avi_1.jgw
20100725142235_trimmed.avi_1.jpg

3.4 *Gidas* - Geospatial Image Database and Analysis System

The Geospatial Image Database and Analysis System *Gidas* was originally developed for the automatic recognition of characteristic sea floor features in organic-rich sediments or submarine mud volcanoes (Lüdtke et al. accepted, Jerosch et al, 2006). Such features are bacteria mats and tube worms, as well as gas related features like small pockmarks in a muddy sea floor or carbonate precipitations. Within the generic framework of *Gidas* the analysis of georeferenced images in general is possible and was tested on Labrador Shelf imagery in this study. *Gidas* is equipped with useful basic functionality, e.g. manual annotation based on a generic classification scheme, GIS file export, access to image metadata, coordinate transformation and basic rendering of maps using the *OpenMap* toolkit (http://www.openmap.org). *Gidas* is still under development (TZI, Germany).

In the following sections 3.4.1 through 3.4.7, every step in the use of *Gidas* is described. The keyboard commands are given in the *Gidas* Guideline section 3.5.

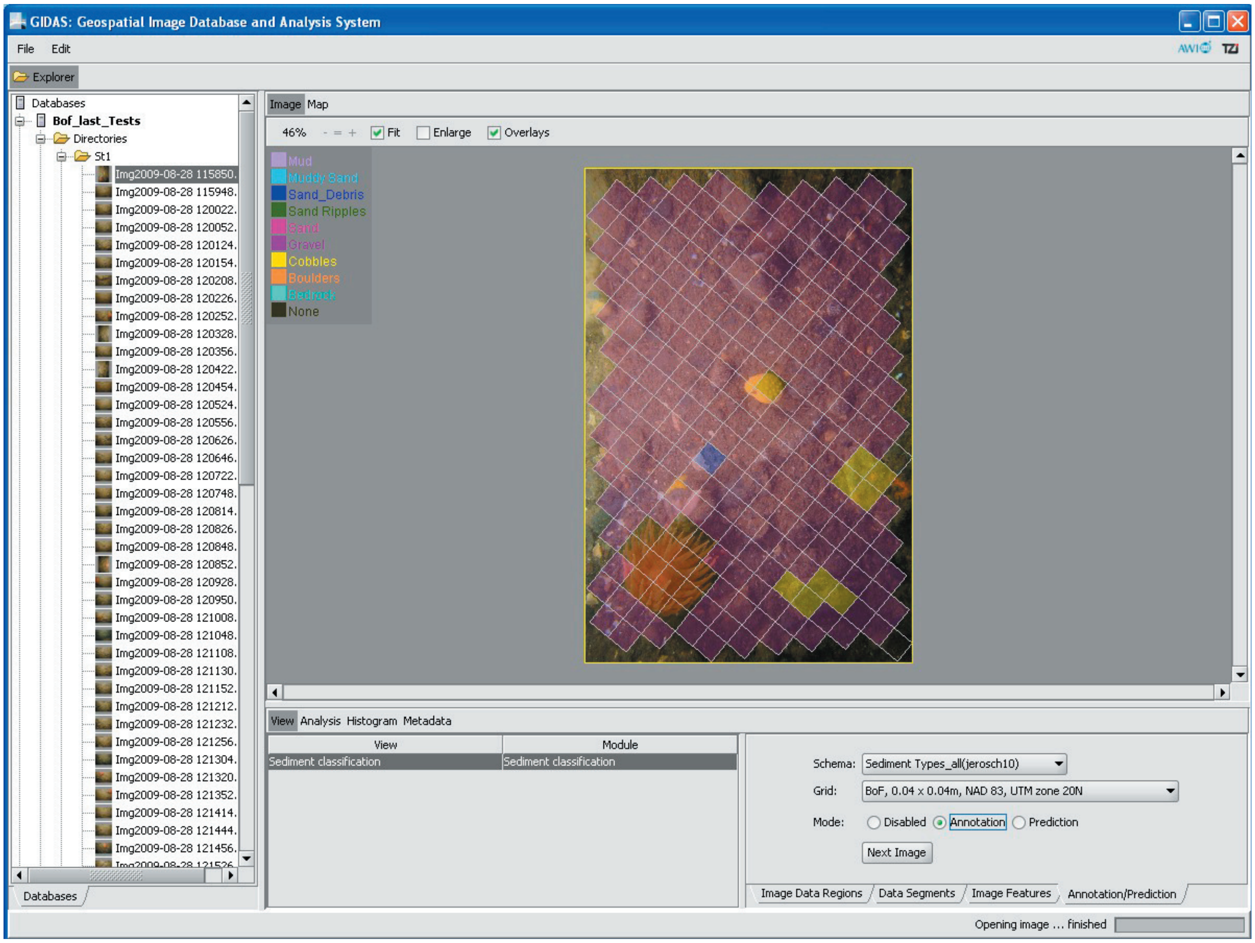


Figure 7 *Gidas* screen shot. A generic grid is the basis for the annotation of sea floor features.

3.4.1 Getting Started

In order for *Gidas* to be used, Java and *postgre SQL* with *PostGIS* and *pgAdmin3* options have to be installed. *pgAdmin3* is used to create databases also used by *Gidas*. Once the database has been created on *pgAdmin3*, it must be initialized through *Gidas* in order to link them both and to import *Gidas* data (tables, modules,...) to the database.

Before the database has been initialized, **schemes** used for annotating (color-based identifying using hexadecimal color codes) must be created as *.xml files and copied into the appropriate *Gidas* folder (F:\...\gidas_runtime\sediment\schema).

Example of a *.xml multiclass **scheme** on sediment texture:

```
<?xml version="1.0" encoding="UTF-8"?>
<schema name="sediments_multiclass (jerosch10)">        scheme name will appear in Gidas
<class name="mud" color="00ff00"/>                      choose class names and colors
<class name="sand" color="d5c90d"/>
<class name="gravel" color="00ffff"/>
<class name="cobbles" color="0000ff"/>
<class name="boulders" color="ffffff"/>
<class name="shells" color="ff0000"/>
<class name="Don't care" color="000000"/>
</schema>
```

Also, before the database has been initialized, **grids** used for annotation (for the detection of certain items more than one grid sizes might be useful) have to be created. The grid name will appear in *Gidas.* Choose the same UTM system as for your imagery data coded as SRID (Spatial Reference Identifier). cx and cy define a location in appropriate UTM coordinates near the center of the investigation area. sx and sy define the grid cell sizes in x- and y- direction in metres (0.07 = 7 cm). The *.xml files must be created and then copied into the appropriate *Gidas* folder (F:\...\gidas_runtime\sediment\grid).

Example of a *.xml **grid** on sediment texture:

```
<?xml version="1.0" encoding="UTF-8"?>
<grid name="BoF, 0.07 x 0.07m, NAD 83, UTM zone 20N" srid="26920" cx="307180" cy="4990388"
  sx="0.07" sy="0.07">
</grid>
```

Note: The easiest way to create the scheme and grid files would be to copy existing files and to modify them respectively.

3.4.2 Data Import and Preparation

Image folders are imported into the database previously created in *pgAdmin3* (see section 3.5.1) on *Gidas* using the correct SRID (Spatial Reference Identifier) following their projection (Note: *Gidas* only works on georeferenced images) using a metric coordinate system such as Universal Transverse Mercator (UTM).

Once importing has completed, the image region of all the photos must be extracted (using the "extract image region" analysis) in order for annotations to occur. If the SRID is incorrectly given when importing, then the image region will not be extracted; if the image region is not extracted, then annotations cannot be made. The necessity of separating areas with data from areas that contain no data when working with video mosaics. If the program is not given the proper georeference (re: SRID), then the grids will not appear on the images and annotating will be impossible. However, two different drop-down menus appear where you can choose the scheme and grid. Annotating can then begin by clicking into the cells and overlaying transparent colors according to the previously defined schemes.

The purpose of using different grid sizes is to detect the sea floor features most accurately. The decision of the best fitting grid can be made by selecting a test data set of images covering all features to be detected and then to run through the whole process for a couple of grid cell sizes. The external analysis tool *Weka* (see section 3.4.4) will tell the user which grid cell size would provide the best prediction results for sediment classes or species in the images of the test data set. This grid size should be used for additional annotations which produce the training data set that will be used for the prediction on the full data set.

3.4.3 Annotation and Feature Export of Training Data Set

Gidas allows the user to annotate image regions based on a georeferenced raster. Users can define the projection and the grid cell size of that raster as well as an appropriate classification scheme for the data set. Then users can label the grid cells and apply the classes of the scheme.

The machine-learning approach relies on the statistics of the provided label information. For example, if starfish are only labeled on reddish sea floors, then the machine learning will not be able to generalize well on other types of sea floor and thus will not be able to detect starfish on, for example, greenish sea floors. The necessary number of labeled cells to specify a class can be tested by the help of cross-validation. In general, the more information that is available, the better the system will be able to generalize to new image data. Ideally, the labels for a specific class should cover all possible appearances for that class. It is recommended to start initially with the test data set of images which were also used to test the best grid cell size (see section 3.4.2) and then continue in order to generate the training data set with even more images selected randomly by using the *Next Image* button (see Figure 4).

Note that annotation should be done as consistently as possible for the best application of the machine learning algorithm. It is strongly recommended to backup the database regularly (see section 3.5.6).

All cells become analyzed within *Gidas* and their texture properties are expressed as 49 texture features such as contrast, minimum and maximum gray level, texture line-likeness and directionality, etc. The annotated cells of the training data set are furthermore assigned to their appropriate classification class and exported as an *.arff file which is a text file with 49 numbers for each cell. Then, *Gidas* uses *Weka* as an external analysis tool to explore the best model of classifiers (e.g., Support Vector Machine, K-nearest neighbors, etc.) to create the training data set. It provides, therefore, the possibility to implement not only one but a collection of machine learning algorithms for data mining tasks.

To determine how many images have been annotated, use *pgAdmin3* and apply the following SQL command to the database: SELECT COUNT(DISTINCT image) FROM sediment_annotated_cell.

3.4.4 Machine Learning Toolkit *Weka*

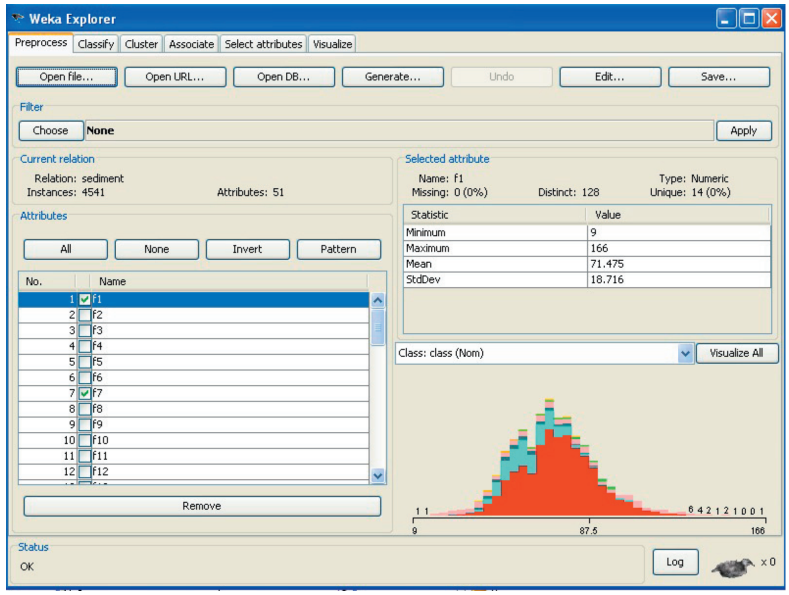


Figure 8 *Weka* screen shot showing the feature distribution.

Weka, developed by Hall et al. (2009), is an open source machine learning toolkit that contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes by applying Support Vector Machines, K-nearest neighbours classifier, C4.5 decision trees, and others (Lüdtke et al., accepted). *Weka* analyzes the similarities of the grid cells using the 49 features and then puts them in the respective scheme class. *Weka* learns the annotations and generates a model (file of type *.model). This model provides quality control by means of crossvalidatation which removes one cell from the training data set and predicts the value of the removed cell. The process is then sequentually applied to each cell. The successful percentage of prediction determines the best classifier which is used to generate the *.model file and later imported into *Gidas*.

The *Detailed Accuracy By Class* provides metrics for evaluating the correctness of a pattern recognition algorithm. **Precision** and **recall** can be seen as extended versions of accuracy, a simple metric that computes the fraction of instances for which the correct result is returned. The set of possible labels for a given instance is divided into two subsets, one of which is considered relevant for the purposes of the metric. Recall is then calculated as the fraction of correct-instances among all instances that actually belong to the relevant subset, while precision is the fraction of correct instances among those that the algorithm believes to belong to the relevant subset.

It is possible to interpret precision and recall not as ratios but as probabilities. Precision is the probability that a (randomly selected) retrieved document is relevant. Recall is the probability that a (randomly selected) relevant document is retrieved in a search.

According to Baldi et al. (2000), the precision for a class is the number of **true positives** (TP) divided by the total number of elements labeled as belonging to the positive class. For example, the number of items correctly labeled as belonging to the positive class (TP) divided by the sum of true positives and **false positives** (FP). FPs are items incorrectly labeled as belonging to the class. Recall in this context is defined as the number of true positives divided by the total number of elements that actually belong to the positive class (i.e., the sum of true positives and false negatives, which are items which were not labeled as belonging to the positive class but should have been).

In information retrieval, a perfect precision score of 1.0 means that every result retrieved by a search was relevant with no indication that all relevant documents were retrieved. Conversely, a perfect recall score of 1.0 means that all relevant documents were retrieved by the search, but this does not indicate how many irrelevant documents were also retrieved.

Often, there is an inverse relationship between precision and recall, where it is possible to increase one at the cost of reducing the other. For example, an information retrieval system (such as a search engine) can often increase its recall by retrieving more documents, at the cost of increasing the number of irrelevant documents retrieved (decreasing precision). Similarly, a classification system for deciding whether or not, say, a fruit is an orange, can achieve high precision by only classifying fruits with the exact right shape and color as oranges, but at the cost of low recall due to the number of false negatives from oranges that did not quite match the specification.

Usually, precision and recall scores are not discussed in isolation. Instead, either of the values for one measure are compared for a fixed level at the other measure (e.g., precision at a recall level of 0.75) or both are combined into a single measure, such as the **F-measure**. The F-measure is the weighted harmonic mean of precision and recall (Baldi et al., 2000) and is the number to be used for the estimation of class and model quality. Learn more about *Weka* at http://www.cs.waikato.ac.nz/ml/weka/.

3.4.5 Model Application on *Gidas*: Annotations and Predictions

The model produced by *Weka* is imported into *Gidas* and is used to predict classes on the annotated images using the correct schema and grid size. At this point the user is able to see where the predictions were made, correct or not, by switching between prediction and annotation screens. The model is then applied to the remaining data set.

3.4.6 GIS Performance of the Results

The data regions (grid cells) of the georeferenced images are extracted as polygons and transformed into world coordinates. The polygons are annotated with both the results of the automatic predicted values (presence or absence of certain features, quantitative analysis) as well as with the annotated values and exported as a GIS layer (shape file) for further processing. A nearest neighbourhood analysis on the cells allows the program to dissolve the huge amount of polygons to bigger multi-part polygons within the GIS. Furthermore, a backup option allows the results (annotations and predictions) and the images to be saved efficiently.

Gidas is still under development through a cooperative arrangement with the Bedford Institute for Oceanography (BIO, Canada), the Center for Computing Technologies/University of Bremen (TZI, Germany), and the Alfred Wegener Institute (AWI, Germany).

3.5 *Gidas* - Guideline

3.5.1 Getting Started

PC and image preparation

- georeference images (see sections 3.2 and 3.3)
- install *Java*, *postgre SQL* including *PostGIS* and *pgAdmin3* options
- copy the *Gidas* software (gidas-runtime folder) to an appropriate location

Note: still photographs should be resized before georeferencing e.g. with *IrfanView* freeware (see section 3.3).

Classification scheme

- go to the *Gidas* folder (F:\...\gidas_runtime\sediment\schema)
- make one or more copies of an existing scheme and modify the new schema individually

Grid definition

- go to the appropriate *Gidas* folder (F:\...\gidas_runtime\sediment\grid)
- make one or more copies of an existing grid and modify the new grids individually

Note: It is recommended to create different schemes and grids to be able to identify the best fitting grid (depending on what features are focussed on; e.g. it would not be possible to identify cobbles applying a 2 cm grid). When the database first is initialized, *Gidas* does not provide the possibility to add new schemes or grids subsequently.

Server/Database generation in *Postgre SQL*

- open *pgAdmin3*
- File - Add Server

  Name *individually* (e.g.'Bay of Fundy')
  Host localhost
  Port 5432
  SSL *keep blank*
  Maintenance DB postgres
  Username postgres
  Password *same as for Postgre SQL installation*
  for the remaining options keep the default

- double click on *New Server* (e.g. 'Bay of Fundy')
- right mouse click on *Databases* - New Database

  Name *individually* (e.g.'BoF1')
  Owner postgres
  Encoding UTF8
  Template template_postgis
  for the remaining options keep the default

- close *pgAdmin3*

Gidas Database Administration

- open start_dbadmin.bat (F:\GIDAS\...\gidas_runtime)
- rider: *Initialize*

  Display name: *individually (*e.g. 'BoF_sediments')
  Host localhost
  Port 5432
  Database name 'Bof1' (as defined in *pgAdmin3*)
  User: postgres

PW: *same as for Postgre SQL installation*

- *Initialize* button (scroll right). Initialization will be finished as soon the button turns white again from gray.
- close *Gidas Database Administration*
- check *pgAdmin3*. Database 'Bay of Fundy'- schemes - public - tables should show defined schemes, grids, classes,... in the distinct tables of the database (table-like button above allows to view the data of the selected table)
- close *pgAdmin3*

3.5.2 Data Import and Preparation

Gidas

- start.bat in F:\...\GIDAS\gidas_runtime (the new database e.g. 'BoF_sediments' should appear automatically)

Import of Images

- right click on Directories, an under-directory of your database (*Import directory tree*)
- choose the folder containing the georeferenced images and the corresponding SRID
  (e.g. 26920 = UTM 20 NAD1983; 26921 = UTM 21 NAD1983)

Note: At that level, all commands (e.g. Batch analyses) are applied to all sub-directories!

Data Region

- right click on Directories (*Batch Analysis*)
- *Extract data regions* (shows image data regions in a yellow frame after the batch analysis)

Extract 49 Features per Cell

- right click on Directories (*Batch Analysis*)
- choose the appropriate grid
- *Extract features per cell*
- repeat this procedure for each grid you want to be tested

3.5.3 Annotation and Feature Export of Training Data Set

Annotation

- click on an image and recognize the four tabs in the lower part of the *Gidas* window:

View	Annotation and Prediction
Analysis	Batch analysis applied to just one image
Histogram	Histogram of the image
Metadata	Metadata of the image

- click *View* and then *Sediment classification* and see four evident tabs

  Image Data Regions
  Data Segments
  Image Features
  Annotation/Prediction

- click *Annotation/Prediction*

- choose both, the annotation scheme and the grid
- start to annotate by clicking on the grid cells, more clicks would change the classification classes

Note: By keeping the mouse pressed and by dragging it, the annotation would extend and therefore accelerate the process significantly. Start with the test data set which should include all classes (see section 3.4.2) and then continue the annotation of more images which are selected randomly by using the *Next Image* button (see Figure 4).

Arff-Export

- to export the annotation joined with the cell features click right on Directories (*Multi Image Analysis*)
- Export training data as ARFF (will create a train.arff file in F:\...\GIDAS\gidas_runtime)

Note: It is recommended to rename the train.arff because *Gidas* would overwrite it applying the next *.arff export.

3.5.4 *Weka*

Weka Classification Model

- got to F:\...\GIDAS\gidas_runtime\lib\modules\sediment
- weka.jar opens the *Weka* GUI
- *Explorer*
- *open file* and navigate to the created *.arff (e.g. train.arff)
- *classify* (next rider)
- Classifier - *Choose* (keep the rest default)
    - choose a classifier, either a *machine learning* function (e.g. SMO, a sequential minimal optimization algorithm for training a support vector classifier applying PolyKernel) or a *decision tree* classifier
    - by clicking on the name of the classifier you get more tuning options for the classifier
- start
- save model (right click on model in result list) as e.g. ‘BoF1.model’

Note: try different classifiers and compare the percentage of correctly classified instances (cells) applying cross-validation to identify the best fitting model.

- SMO with different Kernel, Gamma, try to tune the complexity (higher)
- RandomForest (Decision Trees, vary number of trees), J48 (C45)
- For more information on tuning the classifiers see Lüdtke et al. (accepted)

Example of a classifier output (this summary is given at the end of the *Weka* analysis)

=== Summary ===

Correctly Classified Instances	38	90.4762 %
Incorrectly Classified Instances	4	9.5238 %
Kappa statistic	0.8014	
Mean absolute error	0.2459	
Root mean squared error	0.3276	
Relative absolute error	110.6115 %	
Root relative squared error	100.5808 %	
Total Number of Instances	42	

The *Summary* provides a number of statistical mean values accomplished by a cross-validation for measuring the differences between values predicted by the *Weka* model and the values actually observed.

=== Detailed Accuracy By Class ===

TP	FP	Precision	Recall	F-Measure	Class
0.9	0.031	0.9	0.9	0.9	17
1	0.2	0.9	1	0.947	18
0	0	0	0	0	19
0	0	0	0	0	20
0.5	0	1	0.5	0.667	21
0.91	0.136	0.89	0.91	0.887	(Weighted Avg.)

Note: The numerical Id of class *17* etc. can be identified (e.g. *mud*) in the *sediment_annotation_schema_class* table of the database in *pgAdmin3*.

=== Confusion Matrix ===

a	b	c	d	e	<-- classified as
9	1	0	0	0	a = 17
0	27	0	0	0	b = 18
0	0	0	0	0	c = 19
1	0	0	0	0	d = 20
0	2	0	0	2	e = 21

Note: The numerical Id of class a is *17* and can be identified (e.g. *mud*) in the *sediment_annotation_schema_class* table of the database in *pgAdmin3*.

The *Confusion Matrix* shows the distribution of the predictions, e.g. that 9 of 10 cells have been assigned to class *a*, and 1 to class d.

3.5.5 Model Application on *Gidas*: Annotations and Predictions

Model Import and Application

- applying prediction on just **one** image
  - go to *Analysis* (next to *View*) in the lower area of the *Gidas* window
  - click on *Prediction* and select the appropriate scheme, model (open model manager, import and name model) and grid
  - *run* (execute arrow under the *View* tab)
  - click again on the *View* rider and switch between prediction and annotation
- applying prediction on the **whole data set** (Directory) or **by individual station** (sub-directory; e.g. Stn. 6)
  - go to Directory or sub-directory - right click - *Batch Analysis*
  - click on *Prediction* and select the appropriate scheme, model (open model manager, import a *.model file and specify a model name) and grid

3.5.6 *Gidas* Database Backup

It is highly recommended to backup the databases regularly. This can be done as follows:

Gidas Database Administration

- start_dbadmin.bat in F:\...\GIDAS\gidas_runtime (all existing databases should appear automatically)
- left click on database to select the database to be saved
- press *Backup* button
- choose the folder storing the *Gidas* database copies

3.5.7 GIS Performance of the Results

GIS-Export

- shape file export on the whole data set (Directory) or one station (e.g. folder: Stn.6)
- go to Directory or sub-directory - right click - *Multipart Analysis*
- *Export labels*

Note: At that level the export will store both the annotations and the predictions per cell. Each cell will be a single polygon. A nearest neighbourhood analysis on the cells allow the combination of many polygons to bigger multi-part polygons within the GIS (dissolve command).

The export function can also be used to export annotations before the *Weka*-model is applied in *Gidas*. An export shape file will be created in the F:\...\GIDAS\gidas_runtime folder. It is recommended to rename the export shape file because *Gidas* would overwrite it while applying the next label export.

4 Results for the Labrador Shelf

4.1 Georeferenced Images Extracted From HD Video Footage Displayed in 3D

For this Open File we used video data from station 6 acquired on the Labrador Shelf in 2010 onboard the CCGS *Matthew* 2010–037. The HD camera started to record the seabed in a 3–5 m deep and 20-30 m wide fresh ice scour on the Labrador Shelf (see the red line appearing in Figure 9). The transect consists of 21 min 57 sec of video with an averaged speed over ground of 0.22 metres per second. The movie covers a distance of 285 m and a mean swath width of 2.85 m, therefore, an area of 812 m^2 in a water depth of 134–140 m.

The trimmed video footage and correctly formatted navigation were used to apply the *Video-FrameGeoreferencing-ing-Tool* (section 3.3) and to generate georeferenced frames from the video. The mean pixel size in x- and y-direction in metres (according to UTM 21N (WGS84) coordinates) were calculated as x = y = 0.00220980. A value of 5.2883 was used for the camera rotation (heading). The number of seconds skipped between frames was defined as 12 to avoid an overlap of the georeferenced frames, resulting in a frame frequency of 1 frame / 12 sec. The results in this section are based on 110 georeferenced images extracted from the video footage of station 6.

Zooming into Figure 9 (see Figure 10a) these georeferenced images can be observed on the transect. They are fully georeferenced, meaning each pixel of the image has UTM coordinates. As a result, the images can be draped over the morphology of the sea floor three dimensionally. Zooming in even more (see Figure 10b) enables the recognition of features like rocks in the images, still draped over the seabed.

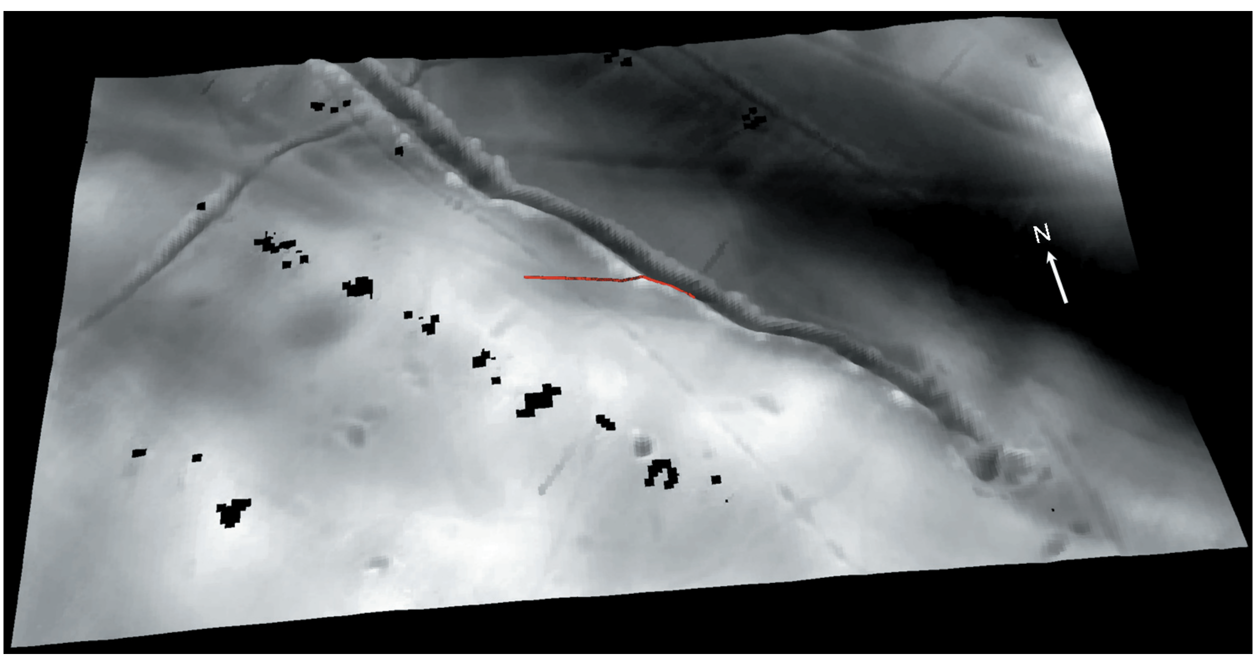


Figure 9 The video transect visualized as a 285 m red line in distance on the 3D multibeam. It crosses a 3–5 m deep and 20–30 m wide fresh ice scour on the Labrador Shelf (see Figure 6).

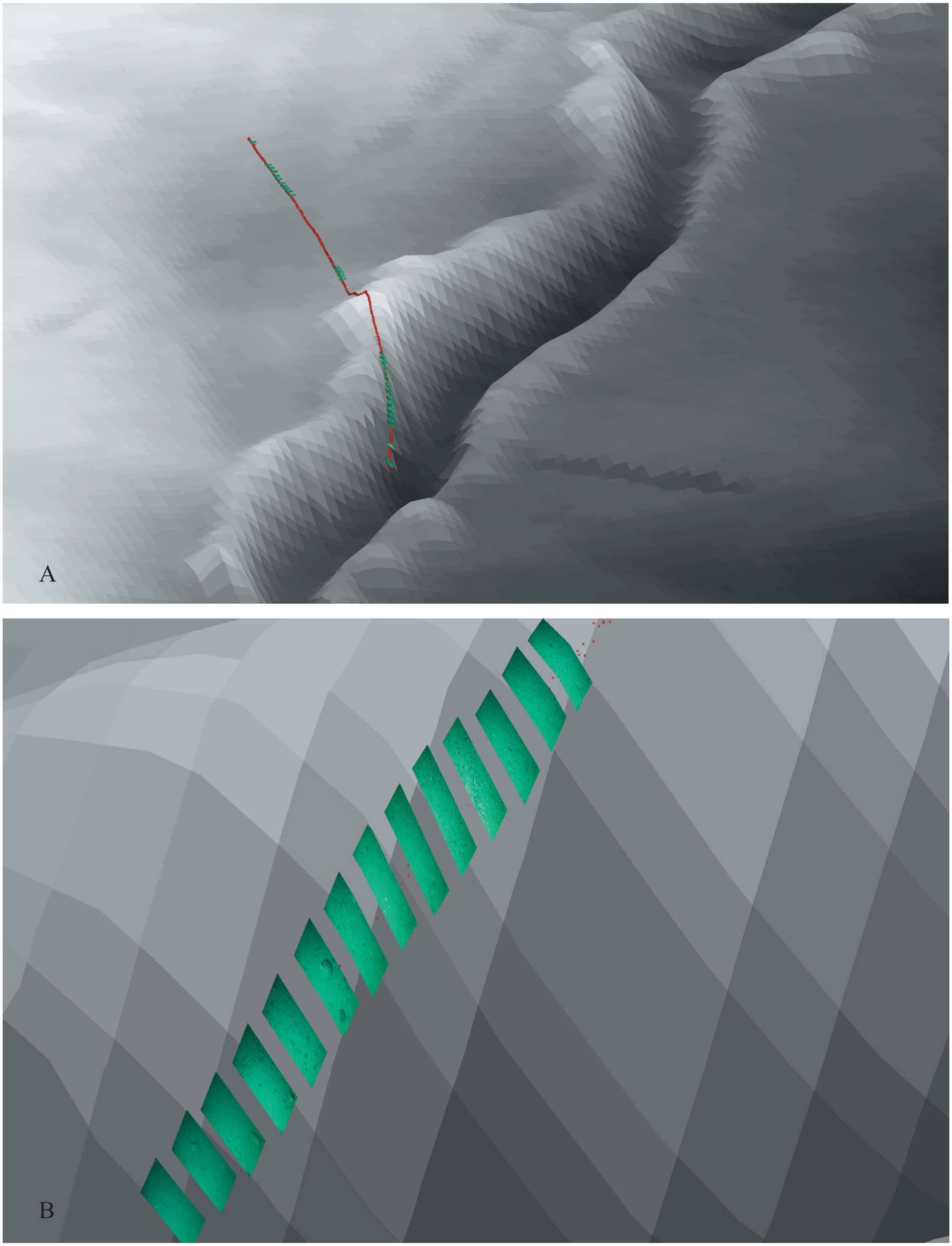


Figure 10 Zooming slightly into (A) so-called georeferenced images can be seen on the transect (navigation shown as red dots). Zooming in even more (B) enables the recognition of features like rocks in the images.

4.2 Automated Detection of Sediment Types on Georeferenced Imagery

4.2.1 Methodological Settings: Classification Scheme and Grid Sizes

The generation of georeferenced images extracted from video footage allows the exact spatial determination of seabed features identified on these images. For further analysis this imagery is imported into *Gidas*. The annotation (Figure 11) is performed by zooming into the images and by manual determination of image areas being either "mud or sand" (green), "granules or pebbles" (light blue), "cobbles" (dark blue), "boulders" (white) and "urchins" on sand (red). Brittle stars on sand have been assigned to be "mud or sand" and other benthic fauna to be ignored (black). The class "don't care" determines negligible observations of other species or blurry image regions.

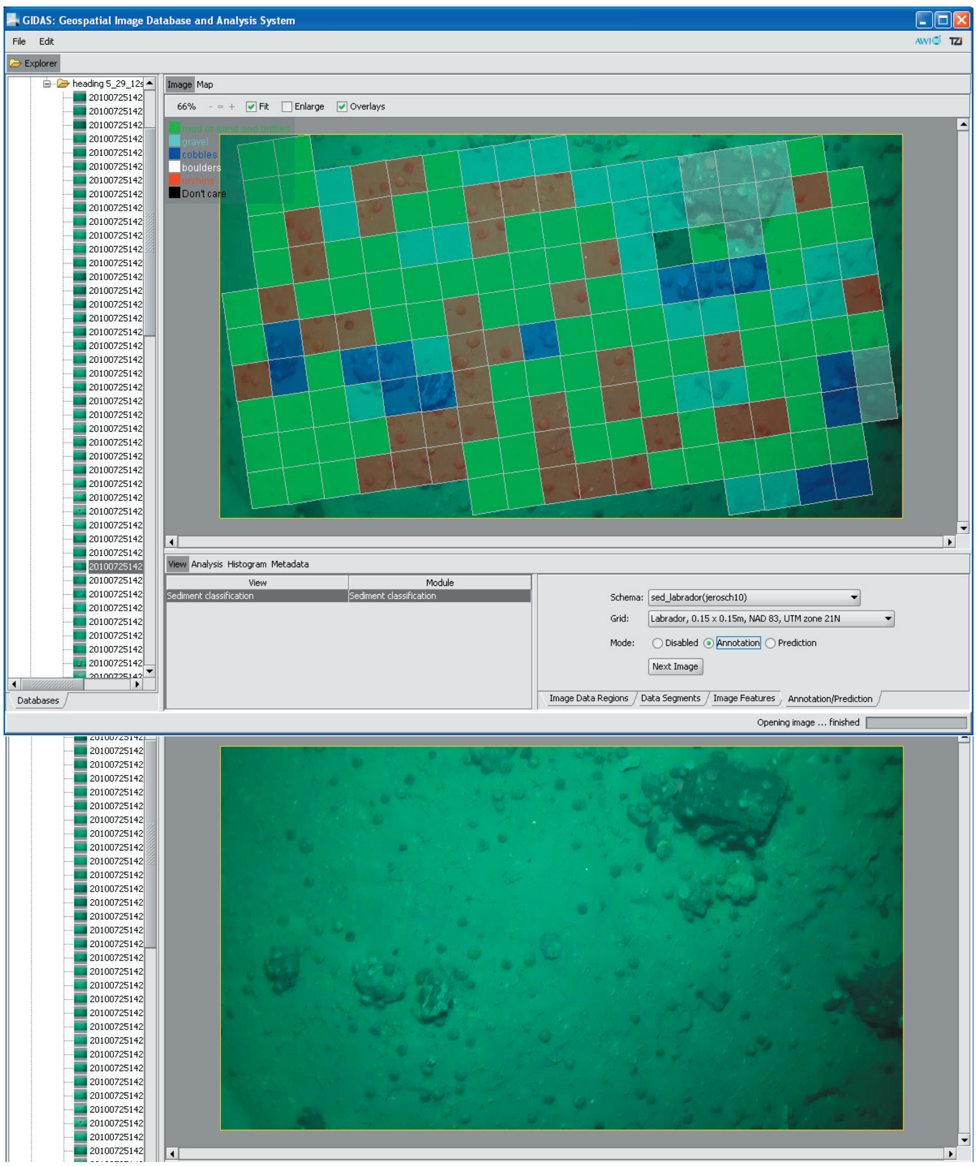


Figure 11 Application of the Wentworth (1922) sediment classification system in *Gidas* (above). Scale is given by the overlay of a 15 x 15 cm grid in UTM 21N (WGS84) coordinates. The lower image shows the non-annotated image.

Grain sized is to be determined after Wentworth (1922). This can be performed with the help of the overlaid metric grid cells projected in UTM 21N (WGS84) coordinates (Figure 11).

In this study “mud and sand” have been combined into one class, therefore, are defined as < 0.2 cm. The term “granules or pebbles” represents a grain size range from 0.2–6.4 cm, “cobbles” 6.4–25.6 cm and “boulders” > 25.6 cm.

In order to test the advantages of different grid cell sized applying the same classification scheme, this study analyses the results of a 7, 15 and 25 cm grid. For the 15 and 25 cm grid, the same 32 images have been fully annotated. In the case of the 7 cm grid just 16 of the 32 images have been used for annotation (Table 1) to avoid a large population of grid cells (instances) and to maintain statistical comparability.

4.2.2 *Weka* - Summary

The annotated cells provide the basis for *Weka* as a collection of machine learning algorithms for data mining tasks. In this study 13,579 cells of the 7 cm grid, 5372 cells of the 15 cm grid, and 1710 cells of the 25 cm grid were annotated. Best results were found by implementing a sequential minimal optimization algorithm (SMO) (Platt, 1998) for training a support vector classifier.

The analysis summary provided by *Weka* supplies a number of frequently-used statistical measures accomplished by cross-validation. Therefore, differences between the values predicted and the values actually observed were measured and serve to aggregate individual differences into single measures of predictive power.

The statistical values in Table 1 consider the mean values of all classes gathered into one analysis and do not allow estimations of single classes. The general trend of the prediction quality is given by absolute numbers and percentages of correct and incorrect classified cells after a cross-validation. Cross-validation is a method to estimate how accurately a predictive model will perform in practice. It predicts a sub-data set of the training data set by using the remaining data set. To reduce variability, multiple iterations of cross-validation are performed using different partitions, and the validation results are averaged over each iteration. Examination of the percentages of correctly classified instances in this study shows that the 7 cm grid is the best size to extract the sediment types from the images, with a value of 76.18 % correctly classified instances.

Table 1 Statistical mean values resulting from the application of a machine learning model to the three (7, 15 and 25 cm grid) training data sets given in the summary of the *Weka* software. The general trend of the prediction quality is given by absolute numbers and percentages of correct and incorrect classified cells after a cross-validation.

	7 cm - 16 images		**15 cm - 32 images**		**25 cm - 32 images**	
Correctly Classified Instances	10,322	76.18%	3850	71.67%	1117	65.3216 %
Incorrectly Classified Instances	3227	23.82%	1522	28.33%	593	34.6784 %
Kappa statistic	0.5074		0.5338		0.4748	
Mean absolute error	0.2316		0.2324		0.2347	
Root mean squared error	0.3249		0.3261		0.3296	
Relative absolute error	129.14%		105.52%		98.928 %	
Root relative squared error	108.53%		98.31%		95.7399 %	
Total Number of Instances	13549		5372		1710	
Total Number of Instances per image	847		168		54	

4.2.3 *Weka* - Detailed Accuracy By Class

Compared to the general investigation of the cross-validation, a more detailed view of the prediction accuracy was performed by means of analyzing statistical measures by class.

F-measure is a measure of a test's accuracy. It considers both the precision and the recall of the test to compute the score. Precision is the number of correct results divided by the number of all returned results and recall is the number of correct results divided by the number of results that should have been returned. The F-measure can be interpreted as a weighted average of the precision and recall, where an F-measure reaches its best value at 1 and worst score at 0.

According to the F-measures in Table 2, the results of this study are reasonable for the classes "mud or sand" and "granules or pebbles". However, the results proved to be poor at identifying "cobbles" and useless for the identification of "boulders", "urchins" and "bedrock". The value for cobbles improves with increasing grid cell size, while the measures for "mud or sand" gets worse.

Table 2 Metrics to evaluate the correctness of a pattern recognition algorithm. Precision can be seen as a measure of exactness or fidelity, whereas recall is a measure of completeness. True positive (TP) is the number of items correctly labeled as belonging to the positive class; false positives (FP) are items incorrectly labeled as belonging to the class. F-measure is a weighted harmonic mean of precision and recall and most suitable for quality estimation and comparison.

ID	Class	TP			FP			Precision			Recall			F-Measure		
		7	15	25	7	15	25	7	15	25	7	15	25	7	15	25
a	mud or sand	0.94	0.91	0.89	0.29	0.19	0.25	0.85	0.80	0.69	0.94	0.91	0.89	0.90	0.85	0.78
b	granules or pebbles	0.72	0.82	0.81	0.17	0.25	0.23	0.55	0.64	0.64	0.72	0.82	0.81	0.62	0.72	0.71
c	cobbles	0	0.18	0.37	0	0.02	0.03	0	0.54	0.53	0	0.18	0.37	0	0.26	0.44
d	boulders	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
e	urchins	0	0.05	0.05	0	0	0.01	0	0.49	0.36	0	0.05	0.05	0	0.09	0.08
f	don't care	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Weighted Average		0.76	0.72	0.65	0.22	0.17	0.18	0.66	0.67	0.58	0.76	0.72	0.65	0.71	0.67	0.59

4.2.4 *Weka* - Confusion Matrix

The issue of prediction accuracy is strongly related to the frequency of occurrences of each class. When a cell in doubt is assigned to the biggest class of the training data set, the probability of a correct classification is larger than if it is assigned to a less frequently occurring class. The probability increases with the percentage of the class in the training data set. Therefore, it is important to know these numbers to be able to interpret the results properly. For a better comparison of the class sizes, Table 3 shows the distribution of the instances assigned to classes for the three models ("Total of Instances" and "% of Total"). The optimal distribution would be evenly distributed instances in order to avoid weighted percentages in the predictions. This problem is reflected by the percentage of "% Wrong" instances that are incorrectly predicted.

The 7 cm grid is obviously not feasible at all to recognize classes other than "mud or sand" and "granules or pebbles". This model achieves the best results for these two classes respectively, but is not reasonable to make any

Table 3 Confusion Matrix (a: mud or sand and brittle stars; b: granules or pebbles; c: cobbles; d: boulders; e: urchins; f: don't care). For a comparison the class sizes are given for each class and grid cell sizes showing the distribution of the instances ("Total of Instances" and "% of Total"). The best distribution would be evenly distributed instances in order to avoid weighted results. Percentages of incorrect predicted instances by class are given in the column "% Wrong".

	Total of Instances			**% of Total**			**% Wrong**			mud or sand (a)			**granules or pebbles (b)**			**cobbles (c)**			**boulders (d)**			**urchins (e)**			**don't care (f)**		
	7	**15**	**25**	**7**	**15**	**25**	**7**	**15**	**25**	**7**	**15**	**25**	**7**	**15**	**25**	**7**	**15**	**25**	**7**	**15**	**25**	**7**	**15**	**25**	**7**	**15**	**25**
a	8642	2441	664	63.78	45.44	38.83	5.59	10.05	10.84	8159	2218	592	483	217	65	0	1	1	0	0	0	0	5	6	0	0	0
b	3008	1855	563	22.2	34.53	32.92	28.09	17.63	19.18	845	297	78	2163	1528	455	0	20	19	0	0	0	0	10	11	0	0	0
c	849	487	158	6.27	9.07	9.24	100	82.55	62.66	199	12	4	650	388	94	0	85	59	0	0	0	0	2	1	0	0	0
d	374	88	28	2.76	1.64	1.64	100	100	100	98	4	4	276	62	10	0	21	14	0	0	0	0	1	0	0	0	0
e	457	375	240	3.37	6.98	14.04	100	94.93	95.42	221	215	149	236	137	75	0	4	5	0	0	0	0	19	11	0	0	0
f	219	126	57	1.62	2.35	3.33	100	100	100	72	39	26	147	57	15	0	28	14	0	0	0	0	2	2	0	0	0
Σ	13549	5372	1710	100	100	100																					

statement about any other class. The prediction results of this model is dealing with two instead of six classes which increases the probability of a correct assignment for them. This has to be taken into account by making a decision about the best model.

The 15 and the 25 cm grid both have their advantages and disadvantages. Both show a better distribution of the classes compared to the 7 cm grid but still have a high accuracy rate (F-measures) in predicting "cobbles", "boulders" and "urchins". The 25 cm grid is the only model reaching an approximative good value for cobbles while the 15 cm grid supplies better predictive values with regard to "mud or sand" and "granules or pebbles".

Furthermore, the confusion matrix in Table 3 shows in which class the wrong instances have been assigned to.

4.2.5 *GIS Performance*

Gidas provides both the annotated and the predicted values as a shape file. The export of the resulting shape file can take several hours depending on the number of images and the grid cell size (number of cells to be exported), because each cell will be exported as a single polygon.

The visualization of the results highlights the differences in the three chosen models. The upper part of Figure 12 shows the annotations on the three training data sets, respectively: manually determined image areas being either "mud or sand" (orange), "granules or pebbles" (blue), "cobbles" (red), "boulders" (yellow), and "urchins" (green). The class "don't care" in black determines negligible observations as other species or blurry image regions.

The lower part of Figure 12 visually presents the predictions applied to the remaining data set. Therefore, the model developed by *Weka* on the basis of the training data sets were imported to *Gidas*. *Gidas* is able to assign the model to the full data set and to generate a shape file of the results (annotations and predictions).

All three models are able to identify sedimentological transition areas from fine sediment ("mud or sand" in orange) to coarse fractions ("gran-

Figure 12 GIS performance of the results.
Upper three panels: Annotations on training data sets: manually determined image areas being either “mud or sand” (orange), “granules or pebbles” (dark blue), “cobbles” (red), “boulders” (yellow), and “urchins” (green). Black color determines negligible (“don‘t care“) and light blue non-annotated images areas.
Lower three panels: Results of the prediction model developed by *Weka* applied to the remaining data set by means of *Gidas*.

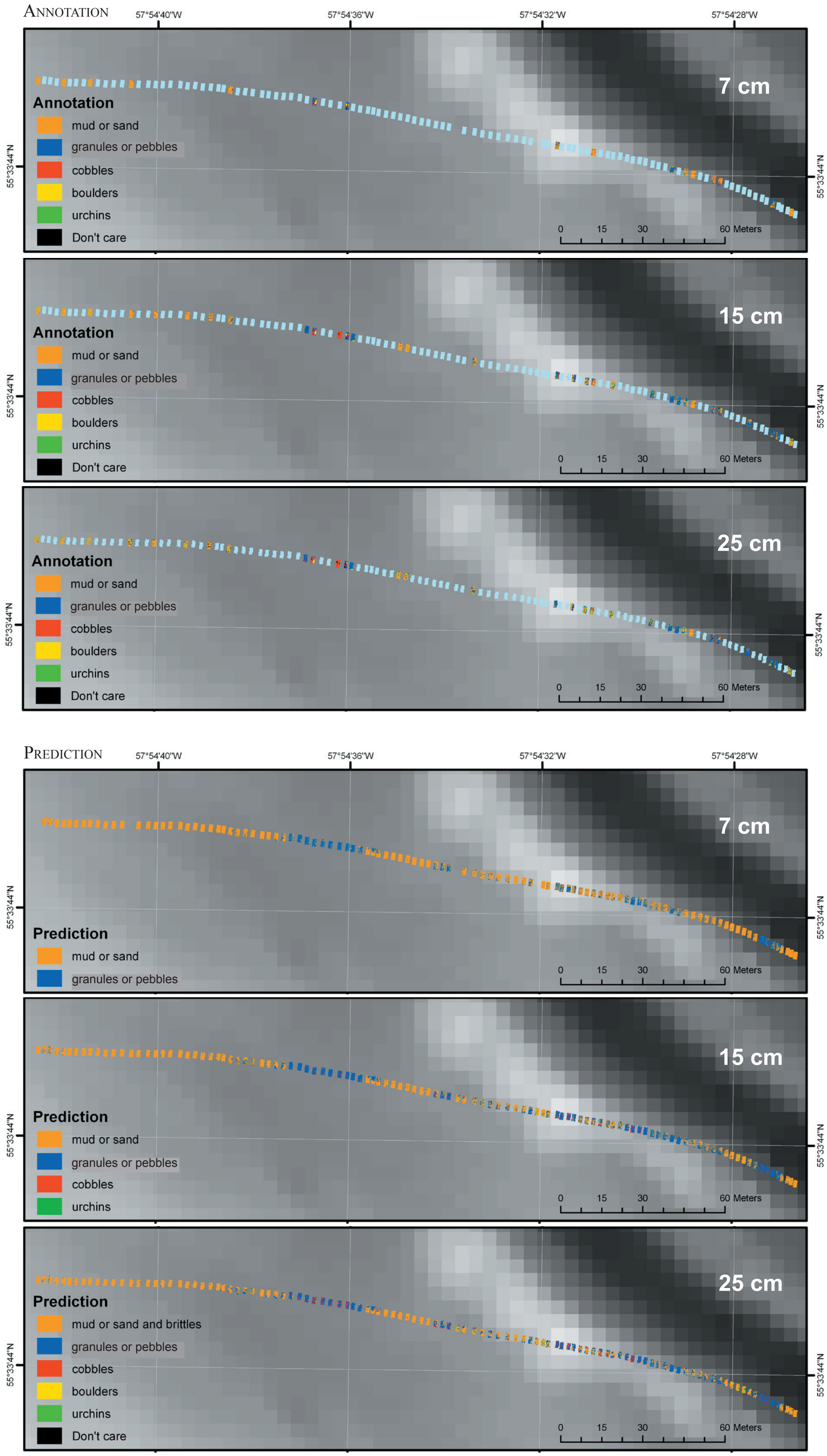
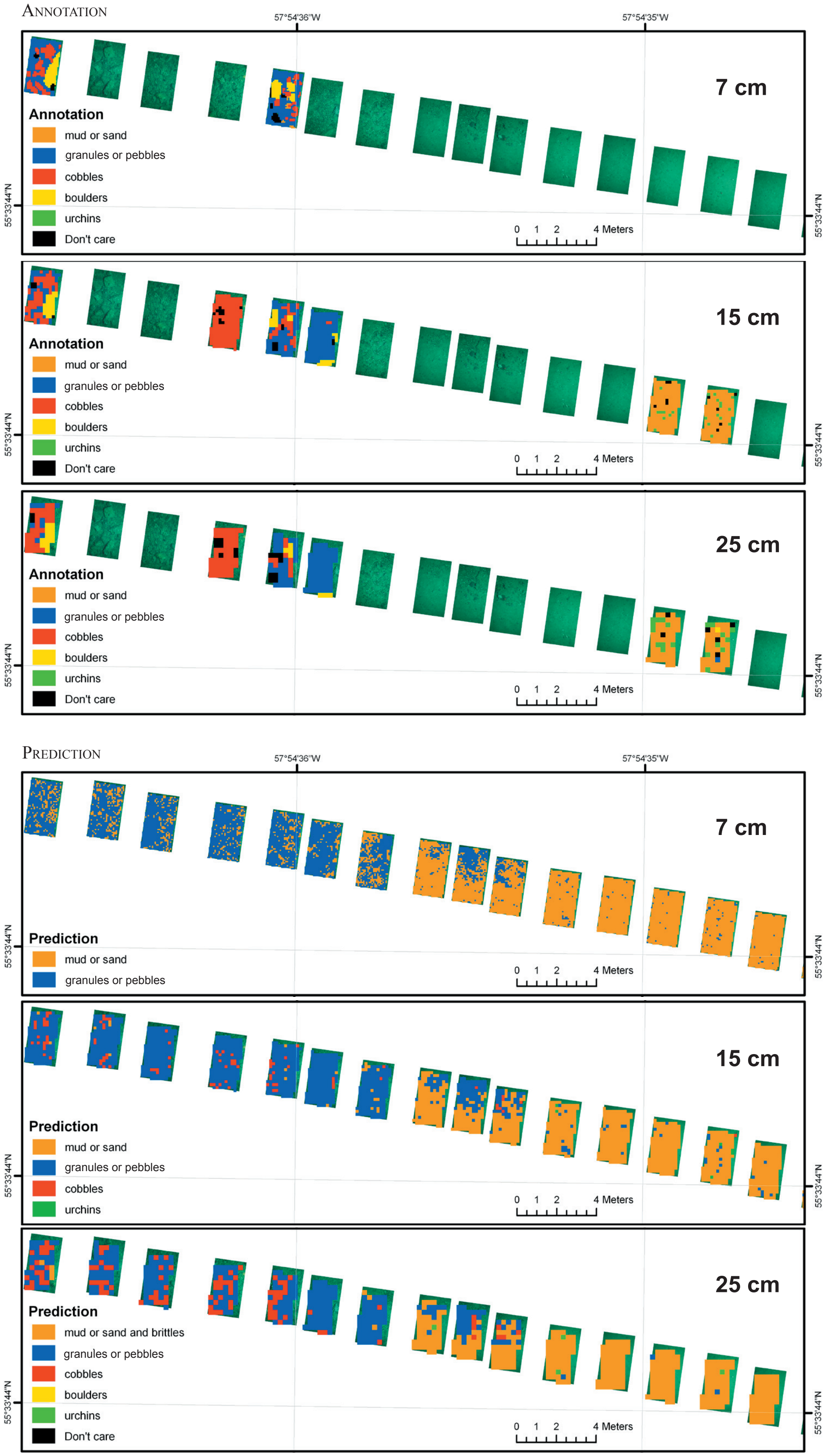


Figure 13 Detailed GIS performance of the results.
Upper three panels: Annotations on training data sets. The manually determined image areas being either "mud or sand" (orange), "granules or pebbles" (blue), "cobbles" (red), "boulders" (yellow), and "urchins" (green). The class "don't care" in black determines negligible image areas.
Lower three panels: Prediction model developed by *Weka* applied to the remaining data set by means of *Gidas* overlain onto georeferenced video frames.



ules or pebbles” in blue). The 7 cm grid gives no information about any additional class, the 15 cm grid predicts some “cobbles” and “urchin” areas, but the confidence in these areas is not high (see *% Wrong* in Tables 3 and *F-measure* in Table 2). This particular prediction inaccuracy is strongly related to the low rate of occurrences of these classes. The only model making a statement on all annotated classes is the 25 cm grid, although the frequency of some classes is so low that they are not recognizable in Figure 12.

Zooming into a section of the transect (see Figure 13), the differences in the annotation as well as in the predictions become more visible. The transition from fine- to coarse-grained sediments are observable even by eye on the non-annotated images. Exemplarily, the 3D-visualization of the 15 cm grid results highlights the correlation between sediment composition and the fresh ice scour on the Labrador Shelf (see Figure 14).

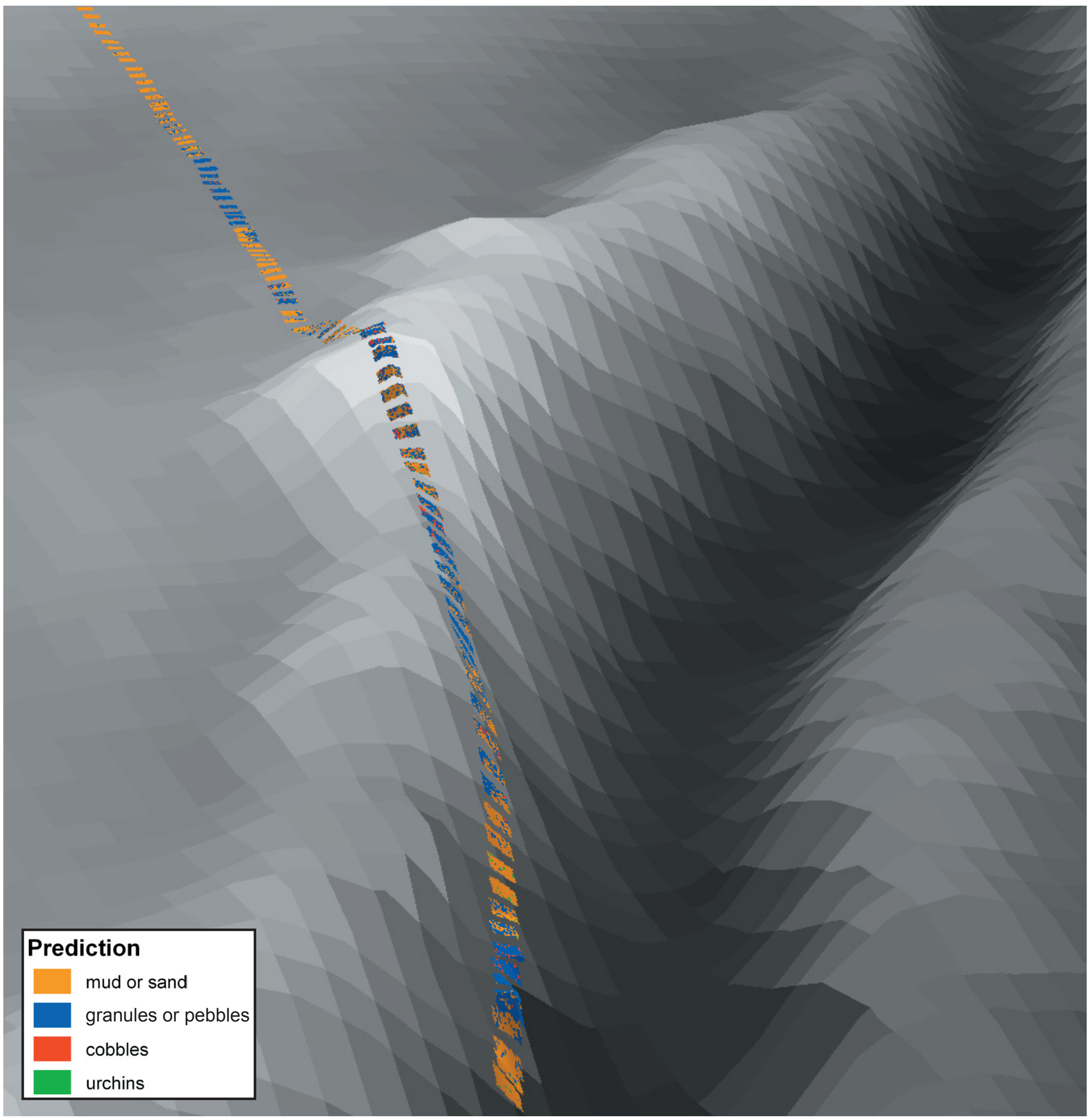


Figure 14 3D-visualization of the 15 cm grid prediction results.

5 Summary and Discussion

This Open File provides a method for gathering and mapping sea floor features from video footage. It provides a guideline through the essential procedure which starts with the video and ends with the visualization of the automatically-extracted image information in a GIS.

To accomplish this, single frames of the footage have to be extracted and then georeferenced automatically by means of the *VideoGeoreferencingTool,* which writes world files for each image. Thus, these images are fully georeferenced and can be visualized in 3D, draped on the contour of the seabed.

The Geospatial Image Database and Analysis System *Gidas* (Lüdtke et al., accepted) is designed to work with georeferenced images. It is a generic software that allows the use to define the classification schemes and the cell sizes used for the annotation of the image features. These cells are part of a georeferenced grid. A set of 49 low-level image features are assigned per cell which describes their visual properties such as contrast, roughness, and texture. Image features are "classical statistical" and "structural textural" attributes which are widely applied to image classification tasks in different domains (Haralick et al., 1973; Haralick, 1979; Tamura et al., 1978; Wu and Chen, 1992).

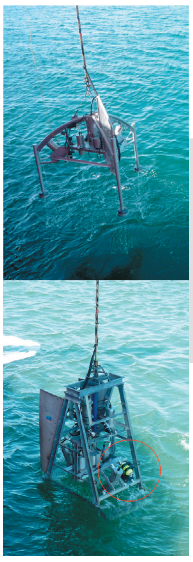
With these features, a prediction model is generated using *Weka* which is a collection of machine learning algorithms for data mining tasks. The quality of the model is assessed by cross-validation. This technique is mainly used in settings where the goal is prediction and testing the accuracy of a predictive model in practice. The best results were found by implementing a sequential minimal optimization algorithm (SMO) (Platt, 1998) for training a support vector classifier.

After the application of the *Weka* model to the remaining data set in *Gidas*, every single grid cell is assigned to a class of the classification scheme. Both, annotations and predictions can be visualized in *Gidas* as well as exported as GIS-readable shape files.

The complete procedure is summarized in Figure 15. The processing path follows the wide gray line which starts with the data acquisition and georeferencing, crosses the analysis framework *Gidas* for annotation purpose, corresponds to the support vector machine in *Weka*, implements *Gidas* again applying the *Weka* model, and ends with the visualization in a GIS.

The quality of the fully automatic detection and quantification of sediment texture with *Gidas* depends strongly on grid cell sizes and the sizes of the features to be detected. All three models, the 7, 15 and 25 cm grid, provide reasonable results for "mud or sand" and "granules or pebbles" and were able to identify transition zones in the sediments from fine- to coarse-grained sediments. The numbers resulting from the cross-validations show that the 25 cm grid would be the best one to include the cobbles into the analysis. An even bigger cell size would possibly improve the results for cobbles, but may be at the expense of the other classes. None of the analyzed grid cell sizes is apparently

Figure 15 on page 31 Summary of the method.

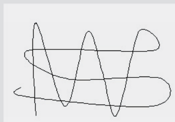


HD Camera mounted onto Campod Heavy (DFO)

Camera system attached to the DFO Campod frame

Navigation

Time stamp and coordinates



Georeferencing

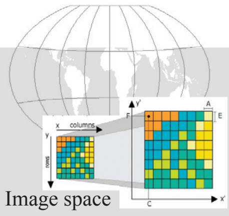
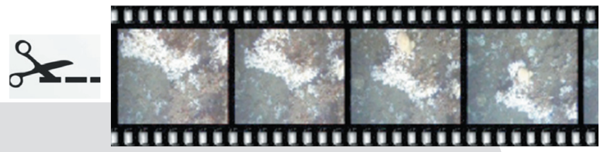


Image space units: pixel

Coordinate space units: meters, feet, others

Raster data is obtained by collecting still photographs and video footage. The location information delivered with them is usually a pair of coordinates for each image coming from a navigation system referred to a time stamp. Thus, to use the raster data sets in conjunction with other spatial data, you may need to georeference them to a map coordinate system. A map coordinate system is defined using a map projection (a method by which the curved surface of the earth is portrayed on a flat surface). Georeferenced raster data may be viewed, queried, and analyzed with other geographic data.

Video Footage

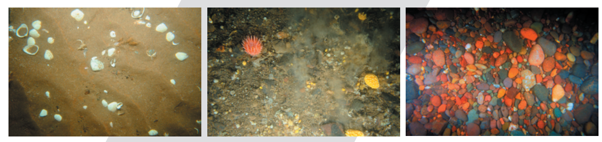


High Definition Camera
- digital (*.m2ts) data
- time stamp extraction using DVMP-Pro-5.0 software
- conversion to uncompressed *.avi using Emicsoft MTS Converter for video mosaicking purpose

Campod Camera
- video tapes
- time stamp on screen
- digitization to uncompressed *.avi

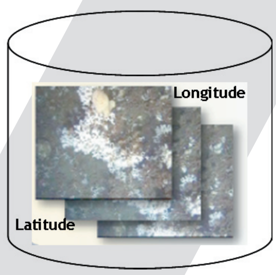
Single frame e.g. every 10 seconds extracted and georeferenced from the video streams (*m2ts or *.avi) and the navigation input using the ***Video-Frame-Georeferencing-Tool*** developed in cooperation with the TZI.

Still photographs



Single still photographs can be georeferenced using the ***Image-Georeferencing-Tool*** developed in cooperation with TZI as long as the geoinformation is stored in the image header (after performing ***GSC NavNet JPEG Merge***,.

Analysis Framework *Gidas*



All images, annotations and predictions, projections,... are stored in a postgis data base analysis in *PgAdminIII* for further analysis in ***Gidas.***

Gidas is equipped with useful basic functionality, e.g. manual annotation based on a generic classification scheme, GIS file export, access to image metadata, coordinate transformation or basic rendering of maps using the *OpenMap* (tm) toolkit (http://www.openmap.org).

Machine learning is a scientific discipline that is concerned with the design and development of algorithms that allow computers to change behavior based on data, such as from sensor data or databases. A major focus of machine learning research is to automatically learn to recognize complex patterns and make intelligent decisions based on data. Hence, machine learning is closely related to fields such as statistics, probability theory, data mining, pattern recognition, artificial intelligence, adaptive control, and theoretical computer science.

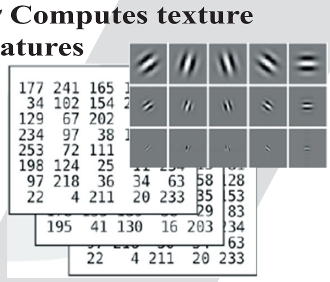
The computational analysis of machine learning algorithms and their performance is a branch of theoretical computer science. Because training sets (annotations) are finite and the future is uncertain, learning theory usually does not yield absolute guarantees of the performance of algorithms. Instead, probabilistic bounds on the performance are quite common.

Expert Annotation within *Gidas*

Gidas allows the user to annotate image regions based on a georeferenced raster. Experts can define the projection, the grid cell size and label the overlaying raster cells as the appropriate scheme class (classification). Also the labels are coming from a generic system which can be modified individually by the expert.

The machine-learning approach relies on the statistics of the provided label information. For example, if starfish are only labeled on reddish sea floors, then the machine will not be able to generalize well on other types of sea floor and thus will not be able to detect starfish on, let's say, greenish sea floors. The necessary number of labeled cells to specify a class can be tested by the help of cross-validation. In general, the more information that is available, the better the system will be able to generalize to new image data. Ideally, the labels for a specific class should cover all possible appearances for that class.

Gidas Computes texture features

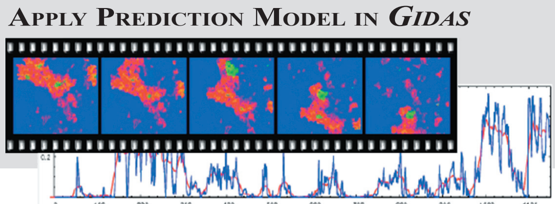


All cells become analyzed within ***Gidas*** and their texture properties are expressed as 49 texture features (as contrast, minimum and maximum gray level, texture line-likeness and directionality,...). The annotated cells of the training data set are furthermore assigned to their appropriate classification class and exported as an *.arff- file (text file with 49 numbers for each cell).

Machine learning in *Weka*

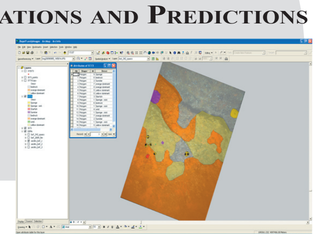
The expert labels expressed as texture features of the 'training data set' are used to generate a quality controlled model applying a machine learning algorithm using an external open-source software (*Weka*).

Apply Prediction Model in *Gidas*



Performing the model on the remaining data set to automatically detect sediment types or sponges in images which have not been assessed before.

Gidas GIS-Tool: Polygon-Export of Annotations and Predictions



Further analysis and habitat mapping by means of a GIS or statistical analysis software.

suited to predict boulders or urchins.

GIDAS is a suitable technique for sedimentological seabed mapping restricted to grain sizes smaller than boulders.

Remarks

- grid cells that are numerous and too small may lead to mistakes in annotation; the users eyes get tired and cause inattentive annotation
- the smaller the cell the better the prediction for small sea floor features and vice versa
- the full image annotation of a certain percentage of the data set provides the frequency distribution of each feature class; this may lead to the awareness that not all classes can be found by automatic detection
- combined results of fully automated predictions and manual annotations (in case of rarely appearing features) can yield significant improvements in the final outcome
- using the results of more than one grid cell size (while detecting the same sea floor features) is not recommended
- the focus on evenly distributed classes regarding the annotation of the training data set would improve the significance of the metrics for evaluating the correctness of a pattern recognition algorithm (Table 2), the confusion matrix (Table 3) and the annotation regarding the frequency of instances by class
- the F-measure (Table 2) is most suitable for quality estimation of the prediction model and for comparison of the single classes
- the confusion matrix in Table 3 indicates to which class the incorrectly-predicted instances have been assigned

Marine research in epibenthic sea floor communities relies strongly on the use of high resolution cameras to analyze the species composition and their correlation to depth and sediments, as well as their spatial distribution. However, the video material currently can not often be used to its full potential as the analysis is time- and labor-intensive and requires the input of taxonomic experts.

Automatic image analysis is mainly based on classic image processing techniques using color, texture, and shape features. Different analysis approaches have different strength. Our ongoing work focusses on testing existing approaches using the video material from the Bay of Fundy where is a good chance in texture recognition for grain sizes and benthic species that cover large areas (e.g., Bryozoa/Hydrozoa mixtures).

6 Acknowledgements

We wish to thank the Natural Sciences and Engineering Research Council Canada (NSERC) who provided support by granting a Visiting Fellowship in Canadian Government Laboratories as a part of the International Polar Year (IPY).

7 References

Baldi, P., Brunak, S., Chauvin, Y., Andersen, C.A.F., Nielsen, H., 2000. Assessing the accuracy of prediction algorithms for classification: an overview. Bioinformatics Review, 16 (5), 412–424.

Barrie, J. V. 1980. Iceberg-scabcd interaction (northern Labrador Sea). Annals of Glaciology, 1, 71–76.

Bass, D.W., Woodworth-Lynas, C., 1988. lcebcrg cratcr marks on the sea floor, Labrador shelf. Marine Geology, 79, 243–260.

Charia, T.R., Peters, G.R., Muthukrishnaiah, K., 1980. Environmental factors affecting iceberg scour estimates. Cold Regions Science and Technology, 1 (3-4), 223–230.

Guinan, J., Grehan, A.J., Dolan, M.F.J., Brown, C., 2009. Quantifying relationships between video observations of cold-water coral cover and sea floor features in Rockall Trough, west of Ireland. Marine Ecology Progress Series, 375, 125–138.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H., 2009. The WEKA Data Mining Software: An Update; SIGKDD Explorations, 11 (1). (http://www.cs.waikato.ac.nz/ml/weka/).

Haralick, R.M. (1979). Statistical and structural approaches to texture. Proceedings of the IEEE 67 (5), 786–804.

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics 3 (6), 610–621.

Harris, I.M., 1974. Iceberg Marks on the Labrador Shelf, Offshore Geology of Eastern Canada. Geological Survey of Canada, 74-30, 1, 97–101.

Jerosch, K., Lüdtke, A., Schlüter, M., Ioannidis, G.T., 2006. Automatic content-based analysis of georeferenced image data: detection of *Beggiatoa* mats in sea floor video mosaics from the Håkon Mosby Mud Volcano, Computers & Geosciences, 33, 202–218.

Kostylev, V.E., Todd, B.J., Fader, G.B.J., Courtney, R.C., Cameron, G.D.M., Pickrill, R.A., 2001. Benthic habitat mapping on the Scotian Shelf based on multibeam bathymetry, surficial geology and sea or photographs. Marine Ecology Progress Series, 219, 121–137.

Lüdtke, A., Jerosch, K., O. Herzog, O., Schlüter, M., accepted. Learning visual sea floor properties for automatic recognition of Pogonophora habitat coverage. Computers & Geosciences.

Platt, J., 1998. Machines using Sequential Minimal Optimization. In: Schoelkopf, B., Burges, C., Smola, A., (Eds.): Advances in Kernel Methods - Support Vector Learning.

Purser, A., Bergmann, M., Lundälv, T., Ontrup, J., Nattkemper, T.W., 2009. Use of machine-learning algorithms for the automated detection of cold-water coral habitats: a pilot study, Marine Ecology Progress Series, 397, 241–251.

Tamura, H., Mori, S., Yamawaki, T., 1978. Texture features corresponding to visual perception. IEEE Transactions on Systems, Man, and Cybernetics 8 (6), 460–473.

Todd, B.J., Lewis, C. F. M., Ryall, P. J. C., 1988. Comparison of trends of iceberg scour marks with iceberg trajectories and evidence of paleocurrent trends on Saglek Bank, northern Labrador Shelf. Canadian Journal of Earth Sciences. 25 (9), 1374–1383.

Wentworth, C.K., 1922. A scale of grade and class terms for clastic-sediments: Journal of Geology, 30, 342–377.

Woodworth-Lynas, C.M.T., Simms, A., Rendell, C.M., 1985. Iceberg grounding and scouring on the Labrador Continental Shelf. Cold Regions Science and Technology, 10 (2), 163–186.

Wu, C.-M., Chen, Y.-C., 1992. Statistical feature matrix for texture analysis. CVGIP: Graphical Models and Image Processing 54 (5), 407–419.