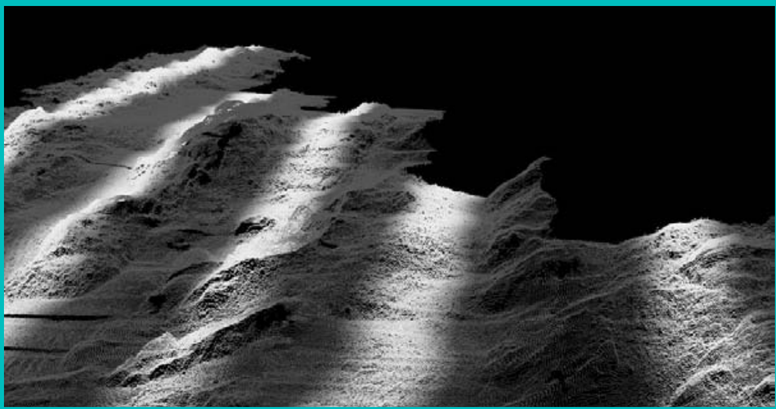
Management of massive point cloud data: wet and dry

P.J.M. van Oosterom, M.G. Vosselman,
Th.A.G.P. van Dijk, M. Uitentuis (Editors)



NCG KNAW

Nederlandse Commissie voor Geodesie Netherlands Geodetic Commission 49

Management of massive point cloud data: wet and dry

Management of massive point cloud data: wet and dry

P.J.M. van Oosterom, M.G. Vosselman,

Th.A.G.P. van Dijk, M. Uitentuis (Editors)

NCG Nederlandse Commissie voor Geodesie Netherlands Geodetic Commission 49

Delft, September 2010

Management of massive point cloud data: wet and dry
P.J.M. van Oosterom, M.G. Vosselman, Th.A.G.P. van Dijk and M. Uitentuis (Editors)
Nederlandse Commissie voor Geodesie, Netherlands Geodetic Commission 49, 2010
ISBN: 978 90 6132 322 8

Published by: NCG, Nederlandse Commissie voor Geodesie, Netherlands Geodetic Commission, Delft, the Netherlands
Printed by: Optima Grafische Communicatie, Optima Graphic Communication, Rotterdam, the Netherlands
Cover illustrations: Shallow Survey 2008, Fugro-Inpark B.V.

NCG, Nederlandse Commissie voor Geodesie, Netherlands Geodetic Commission
P.O. Box 5030, 2600 GA Delft, the Netherlands
T: +31 (0)15 278 28 19
F: +31 (0)15 278 17 75
E: info@ncg.knaw.nl
W: www.ncg.knaw.nl

The NCG, Nederlandse Commissie voor Geodesie, Netherlands Geodetic Commission is part of the Royal Netherlands Academy of Arts and Sciences (KNAW)

Content

Management of massive point cloud data: wet and dry

Editorial

P.J.M. van Oosterom, TU Delft, the Netherlands
M.G. Vosselman, University of Twente/ITC, the Netherlands
Th.A.G.P. van Dijk, Deltares/University of Twente, the Netherlands
M. Uitentuis, ATLIS, Bunnik, the Netherlands

Seminar on 26 November 2009

This publication contains a selection of papers that result of the seminar ‘Management of massive point cloud data: wet and dry’ on Thursday 26 November 2009 at Oracle, De Meern, the Netherlands. This seminar was jointly organized by the subcommissions 'Marine Geodesy' and 'Core Spatial Data' of the Netherlands Geodetic Commission (part of the Royal Netherlands Academy of Arts and Sciences) and the SIM (Spatial Information Management) commission of the OGh (Oracle gebruikersclub Holland).

The theme of the seminar was about the challenges caused by the ever increasing amount of data that is generated by modern sensor systems, both in the wet and dry sectors. To reach a broader audience, both marine (multi-beam echo) and land (LiDAR) data was included. The audience did originate from the Netherlands government/authorities (Netherlands Hydrographic Office, Rijkswaterstaat, Kadaster), research organizations (institutes, universities), and industry (wet and dry data acquisition, Geo-ICT).

It was decided to combine both the wet and dry perspective, with the focus on data management of large point clouds, in order to discover the common challenges to be included in the research agenda. Current solutions may not be sufficient for future needs and therefore new software (data structures, algorithms) and hardware (parallel computing, clusters, grids) need to be investigated.

Prior to the seminar, authors submitted an extended abstract. All 15 presentations as presented are on-line at http://www.gdmc.nl/events/pointclouds. After the seminar the authors were asked to submit full papers and include and reflect the discussions during the seminar. The full papers were then reviewed by the editors and finally the authors submitted revised papers. This publication contains the final selection papers, 10 in total.

Overview of selected papers

The paper by authors of the Hydrographic Office of the Royal Netherlands Navy (NLHO), *Righolt, Schaap, Dorst, and Vos*, focuses on the need for improved and more automated processes, algorithms and visualisation techniques for the validation of massive point cloud data in the Hydrographic field. Large multi beam point clouds may contain erroneous points. The process to find these points is very time consuming. First the authors give an overview of the historic and current use of bathymetric data at the NLHO and the general aspects of data validation. This is followed by an explanation how artefacts in the data are handled. This leads to recommendations for research were both the academic as private sector should cooperate to improve current processing and bathymetric data management solutions.

The combination of both wet and dry point clouds is addressed in the paper by *Kodde* of Fugro-Inpark. An overview of various land-based methods for the acquisition of point cloud data is presented, with some additional attention for Fugro’s own DRIVE-MAP and FLI-MAP systems.

That the point cloud data model is a bit more complicated than just x, y, z values becomes clear when the LAS-format is presented, which allows in addition to store properties such as: intensity, return number, scan direction, classification, user data, r, g, b values, etc. For better dissemination of point cloud data this paper proposes a standardized Web Point Clouds Service (WPCS) in analogy with existing OGC web services. This service should support data streaming and the concept of Level of Details.

The paper by *Swart* describes the properties of the Dutch digital terrain model AHN-2 and its potential use for water management purposes. With some 8 – 10 points per square meter, the AHN-2 is likely to be the most detailed nationwide digital terrain model (DTM). Clearly, this also implies that the data amounts are enormous. The paper explains the need for such high point densities for various tasks of the water boards and the progress that has been made so far in the realisation of this DTM. Profiles drawn across dikes are used for stability analysis as well as for mapping toe lines. Several applications are reported to still work with height images instead of the original point clouds due to a lack of suitable software for handling the massive point clouds.

Simons, Amiri-Simkooei, Siemes and Snellen (TU Delft) describe the recent developments in the processing of multi-beam echo sounder data by focussing on two applications, supported by case study results. Firstly, they emphasise the importance of correcting the MBES measurements for errors caused by the unknown sound speed in the water column (i.e. beam steering and conversions of beam angle and travel-time combinations), which may cause ‘droopies’ or ‘smileys’ in the across track direction. Their method exploits the redundancy of echo soundings in the overlap region between two adjacent swathes, so that no additional sound speed profile measurements are required. Simons et al. present a model that predicts the two-way travel times. By minimising the difference between the measurements and model predictions, optimised depths (or sounds speeds to recalculate the bathymetry) are obtained per ping. A case study of the Meuse near Rotterdam Harbour demonstrates that the artifacts due to unknown velocity profiles are successfully removed from the data (in real time) with an accuracy by which shallow beam-trawl tracks remain visible. Secondly, the variation in backscatter intensity, depending on incidence angle and thereby representing variations in seabed morphology and sediment characteristics, is used in the classification of sea bed sediments. Simons et al. present a new approach that uses backscatter measurements per angle and accounts for the ping-to-ping variability in backscatter intensity. Linear curve-fitting is used to obtain the best model to fit to the backscatter measurements of different sediment types and the difference between measurements and model results are minimised. Acoustic classes are identified statistically (multiple-hypothesis testing). Case studies of the Cleaver Bank, North Sea, Netherlands and a small section of the river Waal show that sediment classification results are in good agreement with the sea bed samples and represent the grain size distributions mapped by geologists.

Ledoux (TU Delft) proposes a new “Triangle” data type to store TINs in spatial databases. In this star-based data structure only the star of each point is stored. The introduction of the paper is focussed on the aspect that computers have great difficulties dealing with very large datasets that exceed the capacity of their main memory. In addition the need to reconstruct surfaces with a triangulated irregular network (TIN) respecting the Delaunay criterion is mentioned. Then an overview is given how TINs are handled in software that is on the market and used in the academic world and their current limitations. Hugo Ledoux introduces the Triangle data type and compares this with the currently used way to manage TINs. The data structure is significantly different from what is usually available in a DBMS. It appears that the Triangle data structure is quite promising in terms of data handling and data space efficiency.

Masry and Schwartzberg of CARIS Ltd describe their system for the management and visualisation of massive point clouds resulting from bathymetric surveys. Point clouds are imported in several open file formats and stored in an octree like structure. The point clouds in the octree leaves can be loaded, processed and put back sequentially. A hierarchical data structure allows visualising arbitrary subsets of the point cloud with 20 – 30 frames per second. Data compression is implemented lossless to maintain the high point location accuracy.

One of the great advantages of massive point clouds is the easy insight in spatial (three-dimensional) structures within the data, especially when visualised in interactive 3D fly-through displays. Because of the extremely large point clouds acquired with state-of-the-art techniques, even modern software and hardware cannot handle these amounts of data well. *De Haan* (TU Delft) highlights several rendering techniques that originate from graphic rendering engines used in games and flight simulators, underpinned by experimental visualisation results with LiDAR (Light Detection And Ranging) data. Real-time visualisation of these amounts of data is a constant play between smart structuring of the data, optimising levels of detail to viewing distances and using efficient ways of paging data in and out of memory. Examples of the AHN2 data set are used to show how point clouds can be used to best extract the terrains, thereby avoiding artefacts such as ‘draping’ and holes. The balance between rendering performance and visual quality is discussed. Future challenges are to develop automated approaches for more flexible data sampling and selection of levels of detail.

For a data producer such as Tele Atlas (part of TomTom) point cloud data offer a number of opportunities to create and maintain their data sets. *Coppens* argues that these opportunities are there for collecting and maintaining the road geometry (better accuracy), road attribution (pavement, lanes, tunnels), road furniture (signs, curbs, trees), and especially for creating 3D models. For various applications, such as personal navigation, car navigation, GIS, Internet LBS, the importance of the above mentioned data components vary. However, in order to make the opportunities realities, there is a number of challenges which will have to be addressed and solved: to keep representations in line with reality (huge amount of data to be processed in timely manner), to have different representations of reality (consistent multiple level of details), to have high positional accuracy (even sufficient for advanced driving assistance systems, ADAS) and to have usable data for applications (not too big, not too detailed).

The paper of *Broen* (Kongsberg Maritime) also addresses the requirement more sophisticated and automated methods of processing, i.e. validating, correcting and analysing the measurements, of the large point cloud data acquired by modern multibeam echo sounders. Broen presents their Seafloor Information System (SIS) of which the Grid-Engine is a core technology that processes and displays the data in real-time. Digital terrain models (DTMs) are built, whereby different resolutions are coupled to certain scales of the terrain maps. The paper includes a full overview, from multibeam data acquisition, including additional parameters, and real-time ping processing via depth processing to displaying and interpreting terrain models. Grid-Engine is successful in the application of multibeam data but was also applied to DTMs of laser data. The near-future release of the Dynamic Grid-Engine will automatically optimise settings in the gridding process and aims at even smarter memory allocations based on predicted vessel heading.

In the final paper, the concept of the Virtual Continuous Model (VCM) is introduced by *Spoelstra* of ATLIS. The VCM technology is intended to support Hydrographic Offices in better managing the ever-growing amount of data produced by multi-beam echo sounders, while at the same time users expect a faster delivery of products (more up-to-date). Using VCM the data is only stored once in the archive of survey data. The users can define their own digital terrain models (DTM), which are all based on the same and shared data, managed in an Oracle spatial database. These models can be defined for various products (including historic data) and different levels of detail, while it always remains possible to trace back the original sources; e.g. for liability.

Closing words

The organizers/editors would like to thank all persons involved in the seminar: first of all the authors of the papers in this publication for their efforts, next the presenters of contributions not included in this publication (but with presentations on-line available), the audience for the active perception and discussion during the seminar, the KNAW/NCG for supporting the seminar and making this publication possible, and last but not least: Oracle, the Netherlands for hosting the seminar. From our side it was a pleasure to be involved in both the seminar and the publication afterwards.

Enjoy reading the contributions in this publication!

Needle in a haystack

R.H. Righolt, J. Schaap, L.L. Dorst, E.M. Vos
Hydrographic Service of the Royal Netherlands Navy

Abstract

The Hydrographic Service of the Royal Netherlands Navy (NLHO) is responsible for charting the Dutch sea areas in Europe and the Caribbean. Products of the NLHO include official Electronic Navigational Charts (ENC's), Paper Charts and other Nautical Publications.
To meet quality requirements for both military and civilian purposes, the Royal Netherlands Navy deploys two survey vessels. HNLMS Snellius (2003) and HNLMS Luymes (2004). Both vessels are equipped with state of the art hydrographic survey systems. Through various causes, large multibeam data files may contain erroneous points.

In this paper we focus on the need to filter the 'needle' from the 'haystack' of multibeam data and how to ensure that only validated data is used in this process. Automated processes should be used where possible to reduce operator workload and maintain consistency. To that end both algorithms and visualization techniques should be improved to deal with noisy point clouds.
It is anticipated that developments in water column imaging, backscatter based bottom classification and artificial aperture sonar data processing will further increase the quantity of data collected. This will require more data storage, more processing power and improved visualization techniques to cope with this flood of data.
In this complex environment with many innovations in the fields of survey, production technology, policy and management it is important for the NLHO to realize its main objective, 'to serve the mariner: Ex Usu Nautae'.

1. Introduction

On behalf of the Netherlands Government, the Hydrographic Service of the Royal Netherlands Navy is the official producer of electronic and paper navigational charts and nautical publications. The products cover the North Sea, the Dutch Antilles, Aruba, Surinam and adjacent ports.
Primary objective is safe navigation for SOLAS-shipping (SOLAS stands for the treaty on Safety of Live at Sea). The products are composed in accordance with rules laid down within the International Maritime Organization (IMO) and International Hydrographic Organization (IHO).

In order to obtain the necessary information, two hydrographic vessels are continuously conducting surveys to register and record sea bottom changes. Not only changes in depth are registered but also obstructions and wrecks. Also data is exchanged with the Directory of Transport and Public works (Rijkswaterstaat), who mainly survey the inland waterways of the Netherlands and hold responsibility for buoyage.
At the shore based Hydrographic Office the collected data together with other relevant information such as buoys, beacons, lights etc. is stored in databases. The content of the databases is the foundation for all nautical products.
In addition to producing charts and nautical publications, the NLHO makes its nautical knowledge and extensive experience available to support Naval operations in the fields of hydrography, oceanography, meteorology, positioning, tides and marine geodesy.

The aim of this paper is to give the reader an insight into the problems that are encountered in the processing of large amounts of data. Hereto we will give a brief introduction on the NLHO and its data now and in the past (section 2). We will describe the validation techniques in use (section 3), possible ways to overcome the various artifacts that may appear in multibeam soundings (section 4) and (future) ways to identify or diminish these errors (section 5). In relation to finding the needle in the big haystack we discuss the current practice at the NLHO and the role that academia and industry could play (section 6).

2. Historic and current use of data

Starting after the fairsheet era, digital storage media of various natures like tape cartridges, floppy disks CDs and the like were used. Although at the time we believed that we were dealing with large quantities of data, present tools have no difficulty dealing with it.
Since the introduction of the MultiBeam Echo Sounder (MBES) the quantity of raw data has increased significantly. In the beginning automatic pruning processes were implemented but it was found that the algorithms were too basic and important data could be lost. Now all data is labeled, evaluated and verified on board then zipped & stored on a portable hard disk for transport to the Hydrographic Office.

Developments in hardware have significantly improved storage devices. Currently all raw data from a survey can be stored on a single Terabyte sized portable hard disk. With two Hydrographic Survey Vessels active in the North Sea, some 35 of these portable hard disks are received by the Hydrographic Office each year. Subsequently all data is binned (5 m x 3 m bin size) and copied to safe storage.

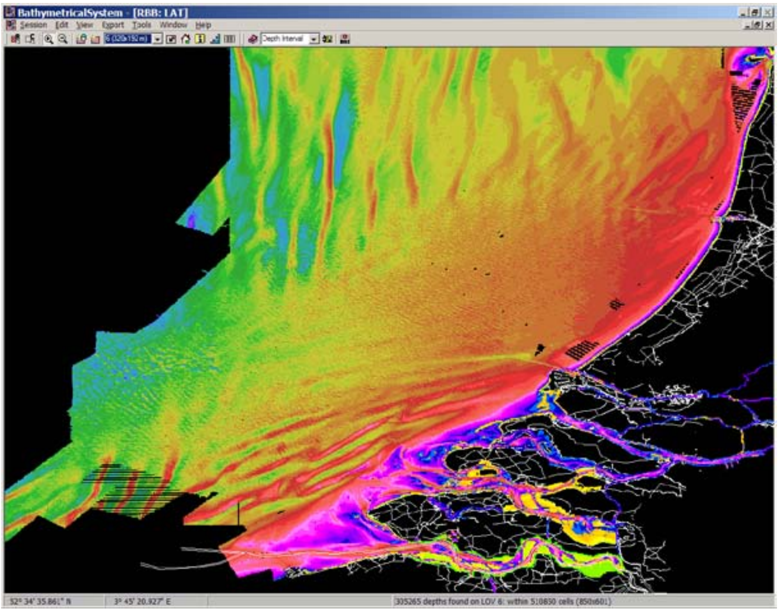


Figure 1. Screen dump of Bathymetric Archive System (BASRBB) on Level of Visualisation (LOV) 6 showing the Southern North Sea, coloured by depth. The area is characterized by sand waves and mega ripples. To better understand the dynamic behaviour of these sand waves studies are underway by various groups at Dutch universities and government institutions.

In the nineties the NLHO developed a software package (Bathymetric Archive System – BAS) that can evaluate, store and retrieve all binned survey data along with its associated meta-data. All data is binned to a worldwide grid, reducing the size of the data files (the cell size is 5 m in latitude by 5 m in longitude at the equator, which equates to 5 m times 3 m at our latitudes.).
A number of tools allow the user to query the database and export selected data in a variety of formats and densities. Due to restraints of the system, we currently only store the Minimum Depth per bin. Other values that describe the quality of a sounding like the Total Propagated Uncertainty (TPU), Standard Deviation (SD), Hits per Bin, Mean & Maximum Depth are used in the evaluation process and subsequently discarded.

Survey data is also received from Rijkswaterstaat (RWS). Their operations are focused on the inland waterways, river deltas and offshore areas out to the 10 m depth contour. Their data is binned to comply with our formats and stored in the BAS database. At present only Minimum Depth and associated meta-data is received. However since the acquisition systems of RWS and the RN navy are essentially the same, it would be possible to devise a storage format that retains all relevant data in a similar fashion.
Within the NLHO these BAS exports are subsequently used by the data processing department to compose and update the bathymetry for the topographical database (TLDB). For the production of Paper Charts we utilize the CARIS Editor, a hydrographic software package that enables us to draw contour lines and select soundings. This package is also used to maintain the TLDB.

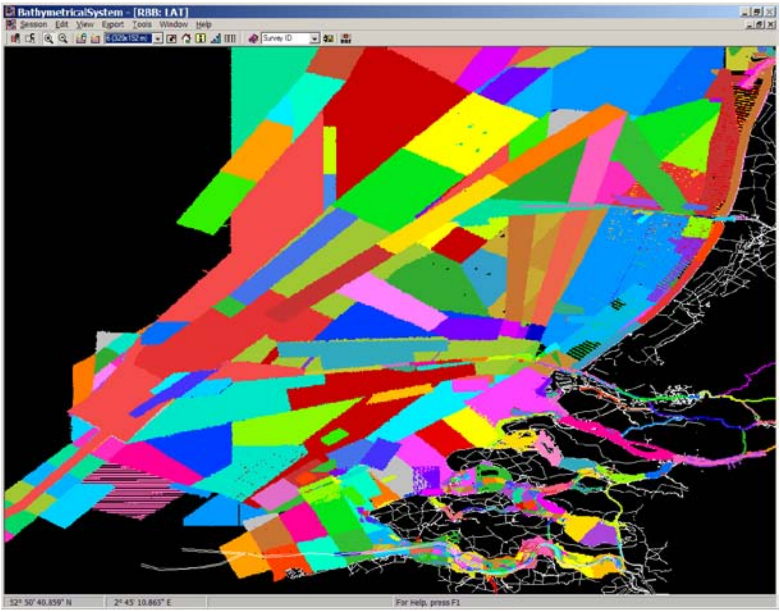


Figure 2. Screen dump of BASRBB LOV6 showing the Southern North Sea, colored by survey. More than 500 surveys are processed by the NLHO each year.

3. General aspects of data validation

The raw data is validated, evaluated and stored. The validation process comprises a number of steps. Its prime purpose is to ensure that the data complies with the accuracy requirements as defined by the IHO Special Publication Nr. 44.
The first step is a comparison with previous surveys of that particular area. In case of discrepancies the cause must be identified. The North Sea bottom is characterized by dynamic bedforms of different spatial scales, each changing at different temporal scales. These changes in the morphology can be explained and are not due to measurement errors. Some other changes may indicate that shoaling has taken place and a dangerous situation has arisen for the mariner. In that case a follow up with a Notice to Mariners (NtM) may be called for.

Secondly, we use the IVS Fledermaus 3D visualization package to inspect and identify artifacts in the point cloud. These artifacts are either introduced by the Multibeam Echo Sounder (MBES) itself or they could be real artificial or natural physical features. In order to be able to recognize these artifacts for what they are, the spatial resolution must be adequate for the intended purpose. A bin size of 1m or smaller is best to achieve this.
Since modern Hydrographic Survey Vessels (HOV) routinely cover large areas, the resulting data files are substantial. A typical survey area of a few square km may contain in excess of 100 million data points. Considering that for every sounding a position, a depth value and some statistically relevant information about the quality of the point must be logged (TPU), the size of

a data files quickly grows to many Gigabytes. Visualizing this survey data as a single image is desirable. It is only when the whole survey is visible, that some of the artifacts become clear. It is then easy to zoom in and pan to an area of interest. For a detailed inspection of the sea bottom full resolution must be used. The IHO S44 rules prescribe that objects of certain size are found and - if possible - identified. At present there is no alternative but to zoom in and pan around the whole work area. The operator must make a judgment as to the nature of the observed features and mark those that warrant further investigation. This is a time consuming, but necessary task!

Manipulating these data files starts to become a problem when time series of the same area are available and need to be retrieved for inspection. Since all data resides on a dedicated office server and are accessed via a local network, this takes some time.
A good visualization allows the analyst to asses the quality of the dataset. Some software will allow the user to toggle between the rendering of the point cloud and the rendering of a Triangulated Irregular Network (TIN) or Digital Elevation Model (DEM) generated from it.
The NLHO utilizes the IVS Fledermaus suite of software. This provides a set of interactive 3D visualization tools for analysis and presentation. It allows us to visually identify and delineate morphologically distinct areas as well as wrecks, pipelines, obstructions and objects. The current version of this software package theoretically supports visualization of up to 100 million points in a rectangular area. Since many survey areas have odd shapes, this means in practice that the point cloud can have far fewer points. Unfortunately the larger surveys need to be split into a number of parts that can be evaluated individually.

The Fledermaus viewer enables the user to rotate the picture and change the viewpoint. It allows the operator to inspect an object from all angles and with a different color palette. Even when viewed at full density and at the perfect viewing angle, the identification of many objects still presents a challenge. Interpretation is highly subjective and depends on operator experience & acquired skills.
For charting purposes it is important to obtain an accurate figure for the safe clearance depth over an object. The nature of the MBES is such that the bottom-track features of the processing software may preclude the proper detection of a protruding object. It is interpreted as a spike in the system because the surrounding data does not confirm it and is edited out. Also the applied weighting parameter should be set to a low value because there maybe very few hits on the shallowest point and averaging can introduce errors. Developments in the area of water-column imaging may alleviate some of these problems (See section 5 Future Developments).

4. Detection of artifacts

Artifacts that were introduced as a result of the MBES and ancillary equipment must be quantified whenever possible. Some of these errors point to a problem in the MBES that can be resolved or minimized by recalibration and careful adjustments of the MBES operating parameters. Many classes of problems cause their own recognizable artifacts. A few are discussed here.
Tidal correction problems manifest themselves as ‘swaths’ that do not match with neighboring tracks. Reprocessing all data with better refined tidal information may be required or in the worst case the survey must be rejected.

Another significant source of errors is the use of an incorrect sound velocity profile. With the use of a Moving Vessel Profiler (MVP) a sound velocity profile is obtained that is subsequently applied to all observations. However due to infrequent sampling and local variations in temperature or salinity, these errors can increase rapidly. Smile and frown artifacts will be the result. With the benefit of hindsight, small improvements can be made by applying the obtained velocity profile retroactively during post processing. Often residual artifacts remain. A software solution that tries to minimize these ‘smilies’ by estimating the best fitting velocity of propagation is

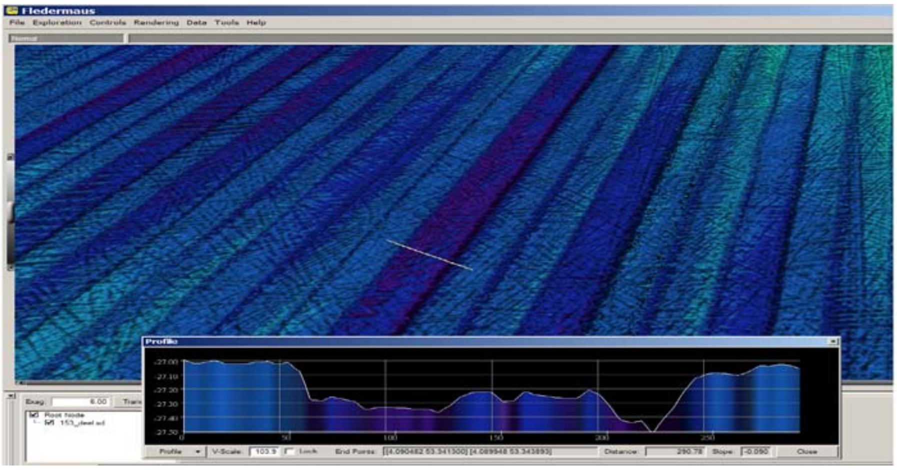


Figure 3. Example of tidal correction error artifacts.

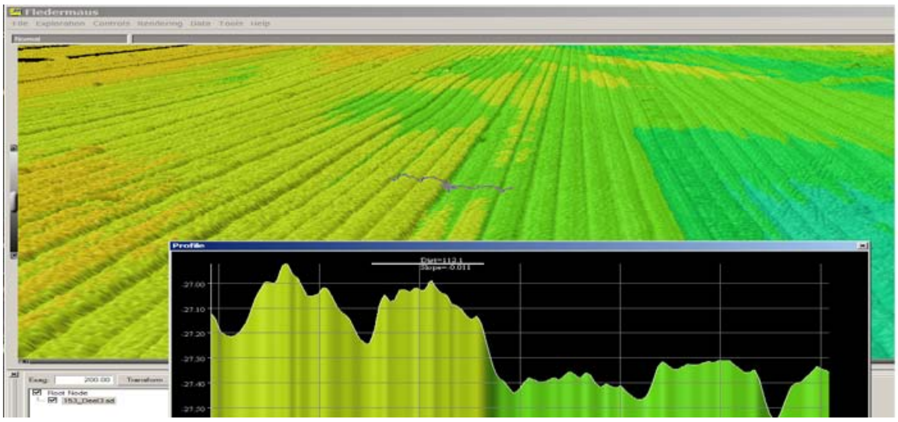


Figure 4. Example of velocity profile error artifacts (smile/frown).

on the horizon but much more work needs to be done to prove its merits before it can be accepted by in the survey industry as a valid tool.

Poor weather conditions may cause heave and roll artifacts to be introduced. The surveyor in charge must decide if the results are still within tolerance. Since many of these errors can be recognized as an artifact, it should be possible to identify the source and subsequently make a correction for it. Software algorithms that operate on the point cloud rather than the swath data can be developed to assist in this process.

Again the IHO S44 document provides guidelines and limits. It is therefore necessary to put numbers to the observed artifacts. Summaries with explanations of error diagrams and examples are available, but often these errors mix and it is not always possible to recognize these artifacts for what they are. It is our wish and hope that some time in the future quality control software will become available that can automatically identify and quantify many of these errors (Hughes Clark, 2003).

Various artifacts that are due to calibration errors or hardware faults may also show up and require attention. Some of these errors can be corrected during post-processing, others must be spotted in the acquisition phase and remedied before the survey commences, such as the misalignment of dual heads of a MBES (Figure 6).

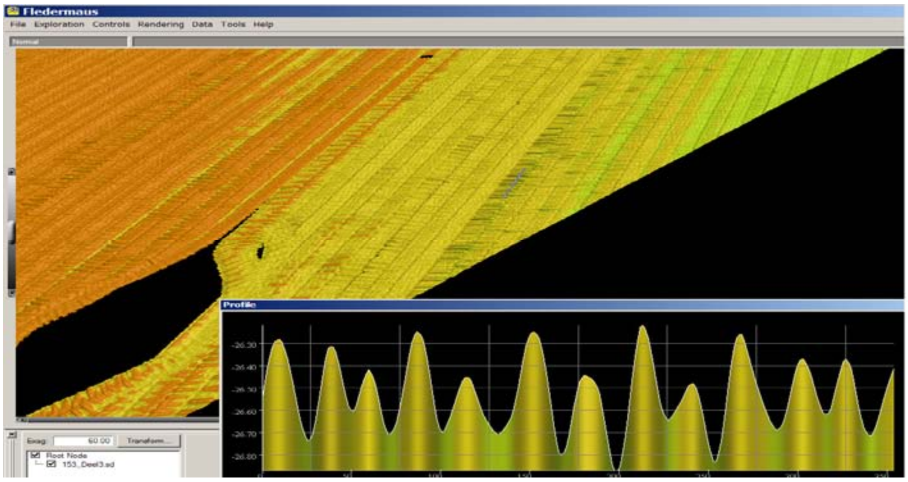


Figure 5. Example of heave artifacts.

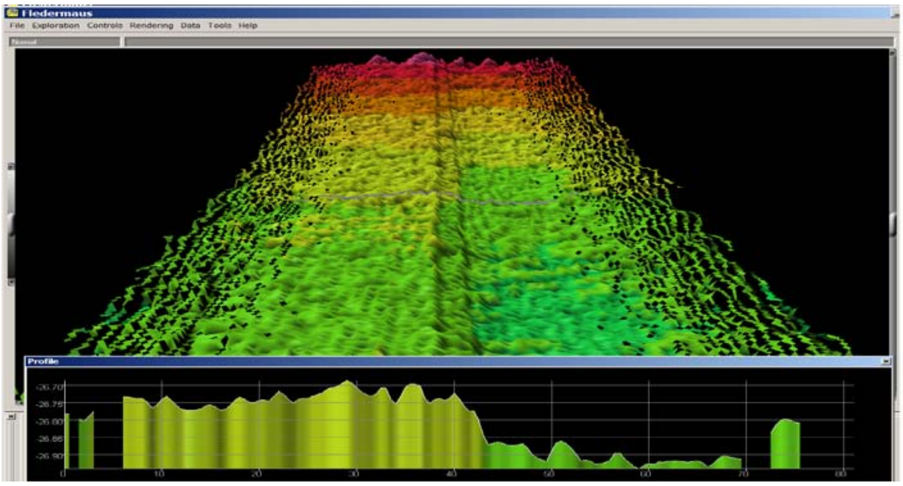


Figure 6. Example of dual-head transducer misalignment error artefact.

5. Future developments

Considering that the current hard- and software can barely meet our need for visualization and manipulation, we have to prepare ourselves for future developments. It is expected that the next generation of acquisition systems will generate an even denser point cloud.

To name just a few systems that are close to reaching acceptance in the survey industry the following developments can be considered.

- The use of multiple Autonomous Underwater Vehicles (AUV) per survey vessel can improve the efficiency of the survey operation with existing technology. No new MBES technology needs to be developed and tested (Kongsberg Huggin – Remus).
- Interferometric sonar systems are close to meeting nautical charting requirements. Advances in electronics and phase deconvolution algorithms promises high-resolution wide-swath bathymetry. Initial results from observations by NOAA indicate that these tools are capable of improving survey efficiency. A swath width of 10 – 15 times water depth seems achievable (Gostnell, 2006).
- Water column imaging is already with us. The data is acquired with the existing MBES equipment, but processed differently. The software to visualize the results still need to be further optimized. Although IVS has published results of their new mid-water visualization tool,

it needs to be proven that it can reliably detect the least-depth of a wreck or object, in which case the need to use valuable boat time for bar-sweep operations is diminished. Possibly algorithms and migration routines may be borrowed from the seismic industry, where 3D-binning and stacking are routine operations. The treatment of data in the seismic industry and water column imaging seem very similar, only the voxel (Volumetric pixel) size is an order of magnitude smaller (Hughes Clark, 2006).
- The Hydrographic Service of the Royal Netherlands Navy is developing a method to extract information about seabed dynamics from bathymetric archives. Software that can make such an intelligent estimate of Seabed dynamics can lead to a review of resurvey frequency. Boat time can be allocated based on need, rather than traffic intensity or perceived dangers. Again such software operates on large quantities of data and hence requires powerful computer hardware (Dorst, 2009).

A treatise on the merits of various visualization techniques and the terminology involved can be found on the internet. A good visualization not only allows the analyst to quickly assess the quality of a dataset, but also enables the planning and control of different processing schemes and ultimately will provide the presentation of the final product.

6. Finding that needle in a haystack

Many survey vessels are active in North Sea and Dutch river estuaries, thus acquiring large quantities of soundings. The Hydrographic Office receives validated soundings from their own surveys vessels and 3rd parties. Of all these millions of soundings only a few thousand can ultimately appear on the charts, making the selection process a critical one.

An automated process that has knowledge of relevant morphological and topographical features can aid this selection process. In this process a massive point cloud need to be searched to identify a small group of soundings that accurately represents the large group whilst emphasizing the nautical hazards. Reduction of the point cloud is currently achieved by a binning process that is implemented as part of the BASRBB software package. Depending on the desired chart compilation scale, this package can further reduce the number of soundings by exporting from a different Level of Visualization (LOV) layer. Every level up will reduce the number of soundings by three quarters. The actual selection is then based solely on minimum depth of the four neighboring bins.

Considering that there are many rules that govern the way in which a sounding is posted there is little room for maneuver once a selection is made (posting a sounding is the process of assigning a sounding to a particular area or feature). Just by inserting an extra sounding one of these rules is easily violated and more corrective action needs to be taken. Our experience with the current generation algorithms is that it is often better and faster to do the sounding selection by hand. It would save time if better algorithms were developed which can take many of these rules into account. Such an algorithm needs to have knowledge of the generalized contour lines, the location of buoys and wrecks, morphological features and a host of other details that determine the posting of a sounding.

Examples of potential hazards are shoals, lost containers, wellheads or wrecks. The mariner needs to be warned about newly identified hazards by issuing a Notice to Mariners (NtM). Note that worldwide up to 10,000 containers are lost at sea in a single year! Due to shifting sands, wrecks can move and tumble over time and thereby possibly changing the minimum safe depth. Whenever a wreck is identified in the point cloud, the minimum depth clearance should be established and deviations from the published values should be investigated further.

We would recommend improving the current practice in the following ways.

Analyzing the remainder of the haystack
To analyze the remainder of the haystack it is essential to be able to distinguish between artifacts and real world features. This pleads for the storage of detailed quality attributes throughout the processing chain.

Error analysis on different levels
Error analysis on different levels of aggregation improves the traceability of both artifacts and real world features hidden in the data. Storage of detailed quality attributes does increase the size of files however.

Visualization
The ability to recognize artifacts depends to a large extent on the kind and quality of the visualization. It is here that improvements can still be made. Also there are algorithms that can remove or mask certain artifacts but by their nature introduce uncertainties with regard to accuracy. In that case it cannot be used for charting purposes, but can still be very helpful to visually identify objects or features that would otherwise remain hidden. For instance, severe heave or roll artifacts can easily obscure features like small wrecks, containers, pipelines, wellheads or mines. The small object detectability as is also described in the IHO S-44 document is thus improved. More investigations should be done to determine appropriate ways for visualizing massive point clouds that represent a marine environment.

In order to reach these goals the knowledge of academia and industry is needed.

7. Conclusion

The existing hard- and software is barely adequate to work with current quantities of data, but new acquisition and processing systems are on the horizon. The expectation is that these new systems will generate ever larger quantities of data that need to be manipulated and evaluated. To that end both algorithms and visualization techniques should be improved to deal with these noisy massive point clouds in an automated way. In this complex environment with many innovations in the fields of computer hardware and visualization techniques, the NLHO is dedicated to be an active player in the field and continues to ‘serve the mariner’: Ex Usu Nautae!

8. References

Brooke Ocean Technology Ltd. Moving Vessel Profiler, MVP-100, www.brook-ocean.com.

CARIS Editor. http://www.caris.com.

Dorst, Leendert L. (2009). Estimating Seafloor Dynamics in the Southern North Sea to improve bathymetric survey planning.

Gostnell, Caleb and Yoos, Jake and Brodet, Steve (2006), Test and Evaluation of Interferometric Sonar Technology. http://www.noaa.gov.

Hughes Clark, John E. (2003). Dynamic Motion Residuals in Swath Sonar Data: Ironing out the Creases, International Hydrographic Review.

Hughes Clark, John E (2006). Multibeam Water Column Imaging: Improved Wreck Least-Depth Determination.

International Hydrographic Organization (2009). IHO Standards for Hydrographic Surveys, Special Publication Nr. 44. http://www.iho.int.

Kongsberg Maritime AS, EM3002D Multibeam. http://www.km.kongsberg.com.

Sang Yun Lee, Kwang-Wu Lee, Ulrich Neuman. Interactive Visualization of Oil Reservoir Data. http://www.usc.edu.

United Kingdom Hydrographic Office UKHO Training Manual, Sounding selection.

The art of collecting and disseminating point clouds

M.P. Kodde
Fugro-Inpark B.V., the Netherlands

Abstract

Point clouds can be collected in various ways, but the most commonly operated sensors include Multi Beam Echo Sounding for sampling water depths and Laser Scanners for assembling point clouds onshore. FLI-MAP and DRIVE-MAP are two examples of dynamic scanning systems, operated from a helicopter and vehicle respectively. A major development is the increasing need for combining data sets from various sources. This requires a common data format to store the point clouds as well as auxiliary imagery and metadata. The current way to do this and new upcoming standards are described in this article. To further simplify the exchange of point clouds, a Web Point Cloud Service is described and proposed. This proposed service allows data vendors and users to operate on remotely hosted point clouds.

1. Introduction

A point cloud is a set of individual 3D points that together represent a surface of an object, terrain, seabed or otherwise. The most common instruments to collect point clouds are laser scanning systems on land and acoustic scanning sensors offshore. Many of these sensors are actually used within Fugro, a world wide operating company that uses a multitude of sensors to capture and model the surface, the subsurface and the atmosphere of the earth, both on sea as well as on land. As such, point clouds have grown to great importance within Fugro over the last years. Developments that contributed to this are Fugro's FLI-MAP LiDAR system, the increasing size of terrestrial laser scanning projects and the advent of Mobile Mapping data using Fugro's DRIVE-MAP system. In addition, the need of combining datasets from various sources is growing.

These developments change the way we work with point cloud data. Existing data models, once developed for airborne applications, are no longer the most optimal system for handling point cloud data from other sources. In addition the typical client base for point cloud data sets is shifting from expert users to a more general audience. In this paper, we present some of the challenges the industry is facing with this shift to new point cloud acquisition systems, and the new demands from the end users. It will be shown that current standards need to be reviewed and that new smart methods for point cloud data distribution are needed.

First we will show some methods for point cloud data acquisition. Then, some examples of the combination of point clouds will be shown, including the combination of bathymetry and lidar data. In the subsequent paragraph, it will be shown why current standards for point cloud data models are not sufficient to store the combined data sets. Finally, a proposal is given on how to improve the way point clouds are used by novel users using web services.

2. Point cloud acquisition methods

On land, Fugro collects point clouds by using various sensors. In general, the selection of a sensor is a trade off between accuracy requirements and the area to cover. This trade off is represented in Figure 1, which shows accuracy and coverage as two dimensions.

The results from tachymetry are usually not called point clouds, since the density of the points is extremely sparse. If a denser point cloud is required over a small area, Terrestrial Laser Scan-

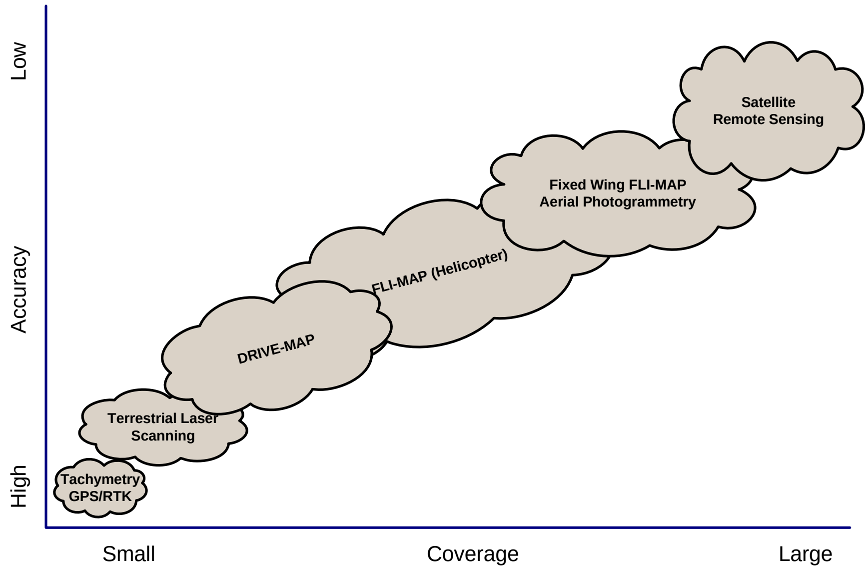


Figure 1. Land based acquisition methods.

ning is the most appropriate method. Terrestrial Laser Scanning is commonly applied in industrial and indoor environments where accuracy requirements are high. Even though a single terrestrial scan may give fewer points compared to dynamic scanning environments, multiple scans are often combined to cover larger areas. Terrestrial Laser scanning can therefore produce very large point clouds as well. This was exemplified by a survey of Fugro on a large industrial site. Up to 150 scans were needed to cover the entire area, resulting in a dataset with more than 2 billion points and a density of 100 x 100 points per m^2.

Larger areas require dynamic scanning systems. Fugro uses its DRIVE-MAP system to cover large highways and urban roads. For DTM generation, the aerial FLI-MAP system is more appropriate (Figure 2). The modern versions of FLI-MAP can be applied from both helicopters and airplanes. Though the carrying system is different, the method of operation for FLI-MAP and DRIVE-MAP are similar in that they both require good GNSS positioning and determination of the attitude using inertial navigation. Both systems acquire a combination of imagery and laser data. The combination of lidar and imagery makes it easier to identify or recognise objects and their attributed information.

Figure 2. FLI-MAP (helicopter based) and DRIVE-MAP.

Near-shore waters and inland waters are typically surveyed with multi beam echo sounding. This method is similar to laser scanning, but measures the run length of acoustic waves instead of laser pulses. The output for both systems, a 3D point cloud, is similar for both systems.

3. Combination of datasets

Often, useful information can be acquired from the combination of datasets. Combined DRIVE-MAP and FLI-MAP data show great advantages in the amount of detail that becomes available for city modelling: FLI-MAP captures the roofs, gardens and the overall DTM, whereas DRIVE-MAP gives a detailed overview of the roads, objects along the roads (such as traffic signs) and the facades of buildings. Figure 3 shows an example of a DRIVE-MAP survey of various buildings. Clearly, a lot of detail is available for the trees, cars and buildings. A combination with FLI-MAP adds information of the roofs and the back of the houses.

Figure 3. DRIVE-MAP data in an urban environment.

Another example of useful combinations of datasets are visualised in the picture below, where a terrestrial laser scan of a bridge is combined with acoustic data of the river below. This dataset, made for a US based bridge, allows engineers to effectively analyse the interface between the bridge and the river currents. An example of this combination is shown in Figure 4.

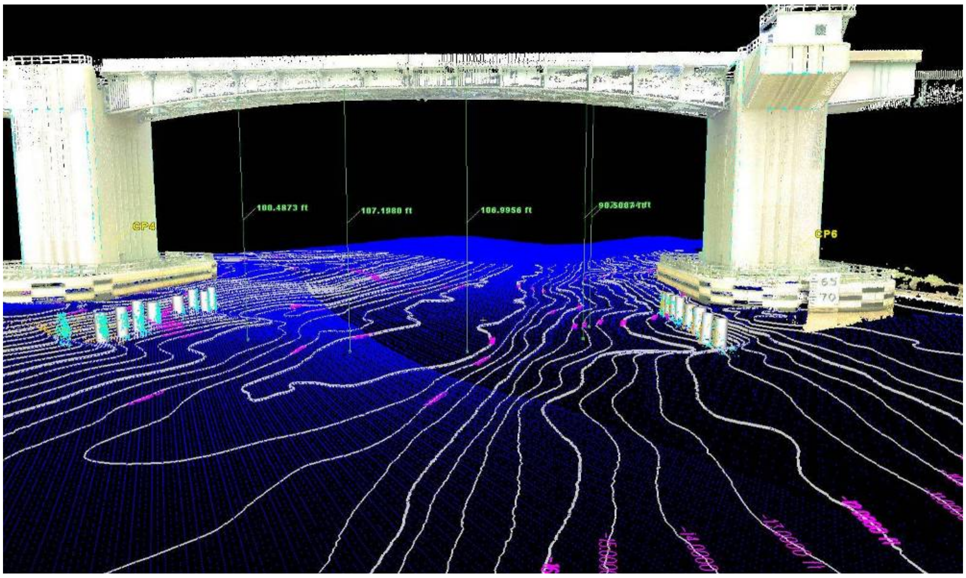


Figure 4. Combination of terrestrial and acoustic data.

4. Point cloud data model

Few data structures could be simpler than a data structure for point clouds. A point cloud is a mere collection of individual points that, together, give a representation of a surface. These points can be stored as a simple list in binary or text format. Therefore, combining multiple datasets is a matter of loading several point clouds at the same time. However, in daily practice, additional information is attributed to these points in order to enhance processing procedures and to store additional information. For the effective combination of datasets, it is important that this information, the data model so to say, remains in tact.

In dynamic laser scanning environments, such as aerial laser scanning or mobile mapping, attribute information is added to each point. For example, each point can be accompanied with the time it was acquired, RGB colour values and classification. Similarly, static scanning applications require knowledge on the scan position for each point and are often attributed with colour as well.

Conventions for adding attributes to point clouds have developed over time, often independently for each type of sensor or even manufacturer. Specifically for airborne laser scanning, a file format was developed to store airborne LiDAR laser points and its attributes. This so-called binary LAS-file, currently maintained by the ASPRS as a standard, allows the attribution of the following properties to each laser point (ASPRS, 2009):

- X, Y, Z,
- Intensity,
- Return number and total number of returns,
- Scan direction, scan angle,
- Edge of Flight line,
- Classification,
- Point source ID,
- User Data,
- R, G, B value.

All mentioned fields are required, except for the intensity value, the user data and the RGB colour value. Most fields are mainly of a technical nature. The values stored in these attributes are directly related to acquisition method of the point cloud: airborne laser scanning. The LAS-standard also gives a list of possible values for the classification field. Each point could be attributed with a certain meaning, to be selected from the list presented in Table 1.

0	Created, never classified
1	Unclassified[1]
2	Ground
3	Low Vegetation
4	Medium Vegetation
5	High Vegetation
6	Building
7	Low Point (noise)
8	Model Key-point (mass point)
9	Water

Table 1. Classification value for lidar points.

Again, these classification values are of a technical nature, derived from the filtering processes of aerial laser scanning and the typical end products.

Since the LAS file is one of the few standards for laser data, the file is commonly used for other types of scanning as well. Most notably, mobile mapping data is regularly stored in LAS files and large terrestrial laser scanning projects can be stored in LAS files as well. However, since the LAS format was not designed for this purpose, there is an inherent loss of data. For instance, Mobile Mapping systems often apply multiple scanners on one vehicle, something that cannot be accommodated in the current LAS data model. In addition, the classification rules need to be made more flexible. Terrestrial and mobile scanning applications have introduced new classes, tunnel infrastructure being the most notable one. The point classification should be seen as the semantic model for laser data. This model may be different for each application: a scan of a tree may require a hierarchical model representing leafs and branches, while a scan of a city may only need a subdivision between “ground” and “not-ground”.

To accommodate these new needs, a new version of the LAS standard, LAS 2.0, was proposed (APSRS, 2007). This new version includes room for additional classification values, such as railroad and powerline. It also provides room to select multiple scan sources and freedom in selection of coordinate datums. However, this proposal did not become an accepted standard and was followed up by the Technical Committee E57 of the ASTM standards organisation. Within the working group for data interoperability a new file format was designed, the ASTM E57 3D file format, or E57 file for short (ASTM, 2009). The file is based upon an XML structure with binary storage of large data volumes such as points and images. The file format adds among many other new features the selection of coordinate systems, point grouping and the combination of 3D point clouds and images. This latter feature, the combination of points and images is very useful, as most modern data acquisitions systems, such as FLI-MAP and DRIVE-MAP, intrinsically combine the acquisition of images and points. Finally, the format contains room for further user defined extensions, to accommodate the specific needs for some sensors and for forward compatibility to new sensors. The E57 file format is not accepted as a standard yet, but it is expected to happen soon.

5. Point cloud webservice

Since computers are getting faster, increasingly the end product of a scanning survey is just the point cloud itself. The common way to use point clouds is and was to derive a data set from the point cloud, e.g. a CAD drawing. Specifically for Terrestrial Laser Scanning, the trend is that modern software allows the end-user to do the modelling or analyses on the fly using specialised software.

Another development is that the usage base of point clouds is growing within organisations. In the past, clients usually were a small group of experts with knowledge of the acquisition technology. New DRIVE-MAP surveys show that the point clouds will increasingly be used by a wider audience, including users with a lot of domain expertise, but little GIS or point cloud processing knowledge. This requires good software to support the user, but also better ways to send, store and maintain the point clouds.

At this moment, USB discs with the point cloud data in a certain format are the most common way to distribute point clouds. However, this method requires large storage facilities at the end user’s IT-infrastructure. A new method for distributing point clouds could resolve this, preferably using web based technologies where the data storage remains at the source. Such a structure would comprise a web based set of standards for distributing, viewing and working with point clouds. This structure basically comprises three components, as shown in Figure 5:

- Point cloud data server,
- Point cloud application server,
- End-user.

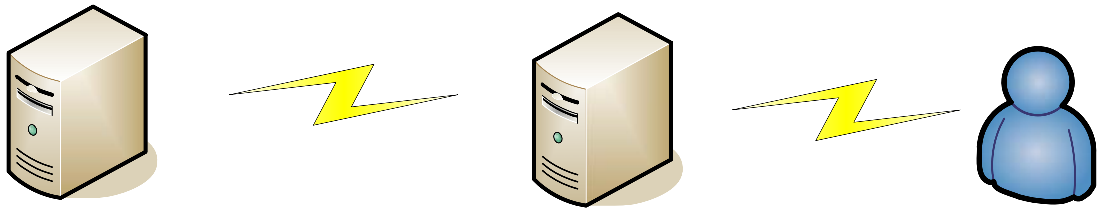


Figure 5. Point cloud distribution structure.

The point cloud data server stores the point cloud data in a standardised data format, such as the E57 file structure described in the previous paragraph. This data should be accessible using, preferably, an open standard. This standard does not exist yet, but in analogy with other OGC standards for distributing 2D GIS maps, it is called Web Point Cloud Service (WPCS) for now. This service provides an interface for:

– Selecting individual points with attributes within a certain 3D bounding box.
– Filtering the point density on the fly, depending on bounding box size and viewing distance.
– Transactional service to update point attributes such as classification, normal vector, etc.

It should be noted that the broadband connection itself is the greatest bottleneck for such a service, since the data volumes are large. This can only be overcome by smart indexing of the point clouds, data compression and the application of Level of Details. Several solutions to indexing and compression exist in the industry, however they are not interoperable.

Similar to the current OGC webservices, the WPCS data streams can be consumed by a program, currently identified as the application server. This program may be a piece of software installed locally on the user's PC, but preferably it is a web based application, allowing the user to do operation from within the browser. There could be a multitude of applications, written in any language such as HTML5/JavaScript, Flash, Silverlight, etc. The only thing they have in common is that they use the same WPCS definitions. Some of these applications, may not even have a user interface, but provide automated processing services only. An example for such a service could be the automated classification of "ground points" in the dataset.

6. Conclusion: point clouds "in the cloud"

This article has shown that various methods for point cloud acquisition exist today. Although these methods vary in there acquisition techniques, they share a very similar end product: a collection of XYZ points. However, the type of information stored alongside these points may be different for each dataset, making the exchange of data between programs and the combination of data sets difficult. It is shown that developments in the standards, including the E57 file format, are providing a solution to this.

Setting standards simplifies the exchange of point clouds, but does not solve the logistical problems of sending point clouds to the end users. In analogy with the current OGC webservices, a Web Point Cloud Service is proposed, to help the dissemination of point clouds over the internet. A smart service should be capable of distributing the data to applications in a bandwidth-efficient manner, thereby opening access to point clouds for many more users. In fact, this correlates with the current developments of "cloud computing", where storage and computing may be scattered through "the cloud", the internet. Bringing point clouds into the cloud may solve many problems and open new perspectives for new point cloud users.

References

American Society for Photogrammetry and Remote Sensing, 2007, LAS 2.0 Format Specification proposal, Madison, AL, US. Retrieved from www.apsrs.org.

American Society for Photogrammetry and Remote Sensing, 2009, LAS 1.3 Format Specification, Maryland, US. Retrieved from www.apsrs.org.

ASTM, 2009, E57 3D File Format Design proposal v2, retrieved from http://www.libe57.org/.

How the Up-to-date Height Model of the Netherlands (AHN) became a massive point data cloud

L.M.Th. Swart *
Het Waterschapshuis, the Netherlands

Abstract

The Up-to-date Height Model of the Netherlands (Actueel Hoogtebestand Nederland, AHN) is a digital terrain model of the Netherlands, owned by the 26 water boards and Rijkswaterstaat. The first generation was completed in 2003. In 2006 the steering committee decided for a second generation AHN with upgraded specifications to comply with the increasing requirements of the water boards. Pilot data was assessed on its suitability for the most demanding application: dike management. In this paper the development to the second generation AHN and its specifications are described. Also the available products and the assessment are described. The massive amount of data and limitations of hard- and software cause the point data to be much fewer used than the grid data.

1. Introduction

Rijkswaterstaat (the executive Directorate General for Public Works and Water Management of the Dutch Government) and the 26 regional water boards are in charge of water management in the Netherlands. They construct, manage and maintain the waterways, dikes, flood defences and structures in order to protect the Netherlands against flooding, they ensure a proper discharge of excess water and an adequate supply of good quality water for all users. For these tasks they need detailed information on the height of the terrain and dikes.

As soon as this seemed technologically viable, laser altimetry was used to produce a country-wide height model. The history of this first generation AHN, completed in 2003, is described in section 2. Meanwhile, technological development of the LiDar technique made it possible to obtain a height model with much higher density. These digital terrain models, for which some water boards undertook separate corridor laser mapping, were good enough to be used for the legally obliged assessment of dike height and stability. In 2006 a pilot was organised to acquire a next generation AHN. The upgraded specifications of this AHN-2 would make it possible to unify the acquisition of height data for both water management and dike management. It was decided that the AHN-2 would be acquired for the whole of the Netherlands during 2008 – 2012. This development is described in section 3.

Meanwhile, dike experts wanted to check whether the specifications of AHN-2 would indeed suffice to fulfil the legal obligations on dike management. Therefore an assessment of the pilot data for dike management was done. It is described in section 4. The upgraded specifications of AHN-2 cause the amount of data to increase dramatically. In section 5 the products and their usage by water boards are described. Special attention is paid to the point cloud data. Section 6 lists some concluding remarks.

2. History of the AHN

In 1996, the new remote sensing technique of laser altimetry promised to have reached the level of development necessary to be used in business processes. Rijkswaterstaat, the water boards and the provinces started an initiative to cooperate in obtaining a country-wide height model primarily suitable for water management: the *Up-to-date Height Model of the Netherlands* or *Actueel Hoogtebestand Nederland* (AHN). The availability of a high-density height model seemed promising compared to the old height information TOPhoogteMD, consisting of 20 to

50 year old terrestrial levelling data with a density of no more than one measurement per hectare.

However, the LiDar technique (in the Netherlands known as laser altimetry) turned out to be on a lower level of technological readiness than expected. In close cooperation with airborne laser scanning companies, the former Survey Department (Meetkundige Dienst) of Rijkswaterstaat developed a high level of scientific and executive expertise in acquiring and processing the data. Issues covered were GPS and the artefacts it could cause in the data, the influence of the behaviour of the inertial navigation systems (INS), the adjustment of overlapping scanning strips and an extensive treatment of the four error components leading to a detailed description of the error structure of the resulting height model (Brand et al., 2003).

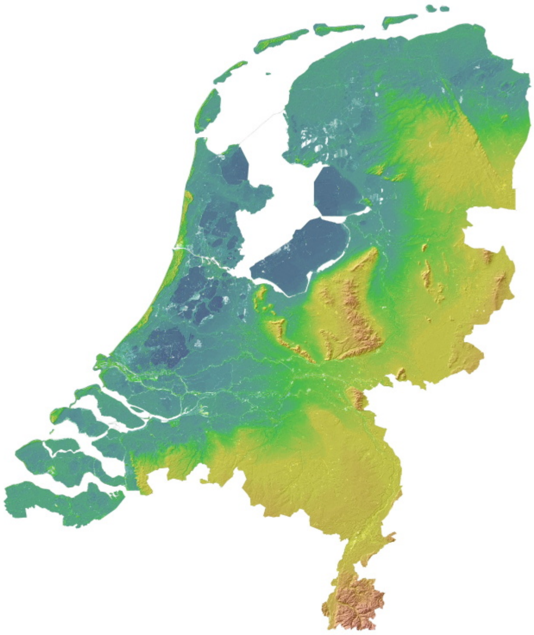


Figure 1. Overview of the first generation of the AHN, completed in 2003.

In 2003 the first generation of the Dutch height model, now denoted AHN-1, was complete, covering the Netherlands with a density of one height measurement per 1 to 16 square metre. As Figures 2 and 3 show, the increasing density is primarily a consequence of technological development and viability and hence almost only depending on the date of acquisition. In a simplification beyond the error structure mentioned above, the height error can be summarized to be at most 5 cm systematic plus 15 cm standard deviation for the stochastic component, for areas not covered with vegetation (Van Heerd et al., 2000).

For many applications, these specifications suffice. Even with the advent of the upgraded second generation of the AHN, the AHN-1 can be of use. Although it is not very up-to-date (cf. Figure 2), it is the only dataset covering the whole of the Netherlands. For large-area, low-detail applications its specifications can be sufficient, e.g. for some water management applications and the study of the topographic features, as the trivial example in Figure 4 illustrates. For some users the much smaller data files and low cost can also be an advantage.

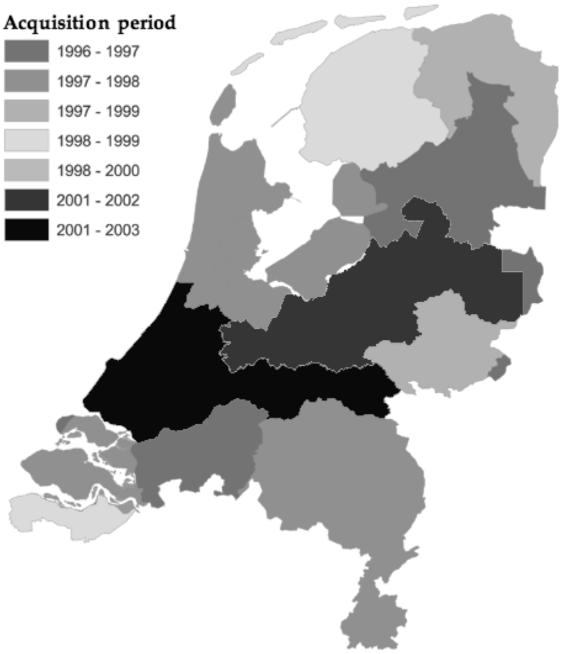


Figure 2. Year of acquisition of AHN-1.

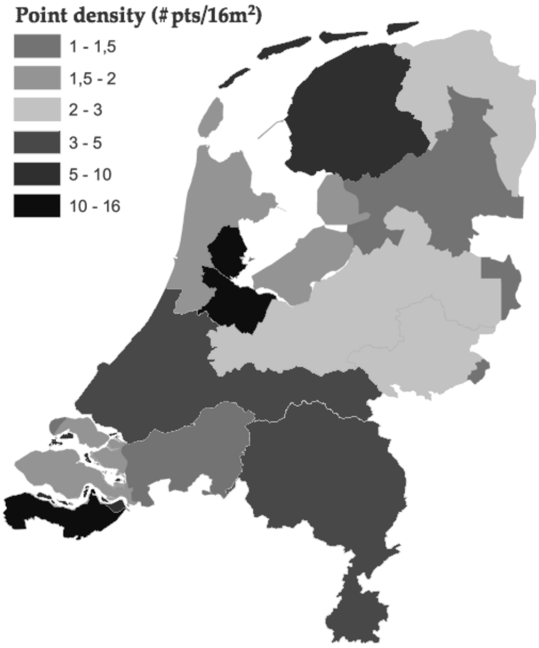


Figure 3. Point density of AHN-1.

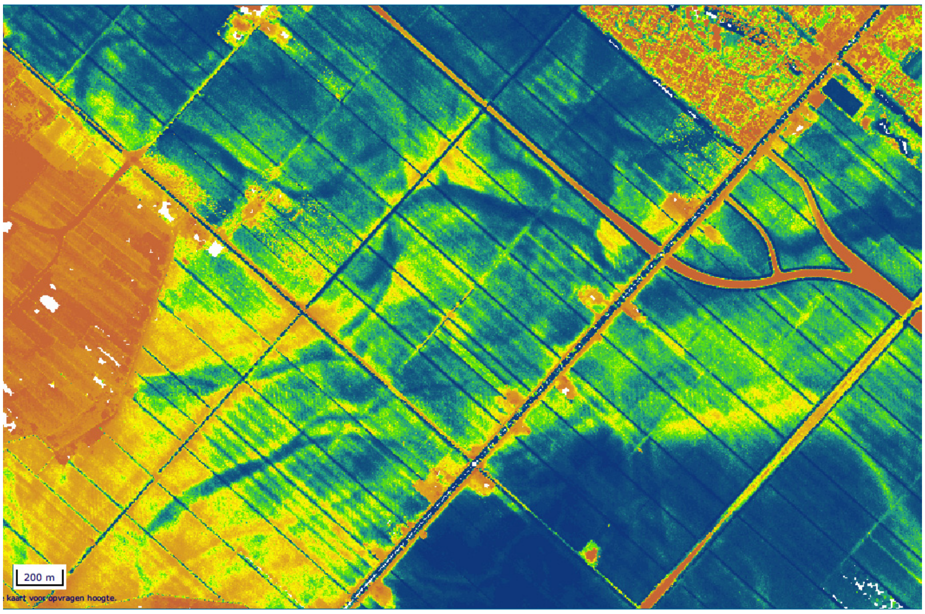


Figure 4. For many applications, the point density and accuracy of the AHN-1 suffices. With the AHN viewer at ahn.nl, any user can explore height data in sufficient detail to discover unexpected features in the landscape himself. Just by adapting the legend of the viewer at the world wide web, the author 'discovered' that the 180 sq km Haarlemmermeer polder he lives in is not as flat as it seems. The 1997 1pt/8m^2 data is sufficient to uncover old streams.

3. Need for a next generation: AHN-2

3.1. Need for higher specifications: dike management

Water boards and Rijkswaterstaat need detailed and precise height data of their dikes. The Flood Defences Act oblige them to assess their dikes and flood defences once per five year in accordance with the very detailed *Directions for the legal assessment of dikes* (*Voorschrift Toetsen op Veiligheid*, VTV) (Van den Berg et al., 2004). In these directions, the height of the top of the dike is important (because of water overflow and wave overtopping), but also its profile, because in combination with the composition of the dike body, this determines the strength of the dike and its resistance against sliding. Furthermore, apart from these primary dikes, also detailed geometrical information of the regional dikes and embankments is needed, because they are considered more and more to be important with respect to flood risk management. Geometrical information is also needed to draw up the obligatory public register of dikes, waterways and flood defences (legger en beheerregister).

Until recently, geometrical data of dikes was mainly acquired by terrestric surveying. The technological development in laser altimetry made it, however, possible to acquire digital terrain models of dikes by means of laser altimetry. As of 1999, some water boards in the large river area had their dikes acquired with LiDar from helicopters. These acquisitions were quite different from the AHN: the point density was much higher, the precision slightly better, simultaneous aerial photography and sometimes videos were available and the mapping was done in strips following the bends of the dikes ('corridor mapping').

3.2. A next generation AHN: pilot with upgraded specifications

Some water boards, experienced in using corridor LiDar mapping with high specifications for dike assessment, considered acquiring a full-area LiDar height model that also could serve their needs on dike management. Private companies were, in the mean time, offering such a high-precision high-density LiDar height model and even had plans to cover the whole of the Netherlands with such a commercially available height model.

In 2006, the AHN steering committee decided to organise a pilot with Waterschap Zeeuwse Eilanden (water board WZE) to acquire a next generation AHN. Its higher specification would make it suitable for dike management, so that separate corridor LiDar mapping with high specifications would no longer be necessary. Meanwhile, a new way of specification and tendering could be tested, leaving more initiative to the market. Also new was the tendering of the quality control to a separate contractor. All contractors had to prove the quality themselves.

The requirements were defined in terms of user needs instead of by detailed technical specifications. Therefore an inventory of current specifications and user requirements was executed and the results were assessed geodetically. An example of specification in terms of user needs is that the requirement on planimetric accuracy was that it should be possible to map objects of 2 x 2 meter with an error of at most 50 cm. The contractor is free to choose a combination of point density, point distribution and systematic and stochastic planimetric accuracy that leads to fulfilling this requirement (the relation between them was treated by Vosselman, 2008b). Also the operational procedures are fully left to the contractor.

Apart from the planimetric requirement mentioned above, the height model should have an accuracy in height of at most 5 cm standard deviation and 5 cm systematic error. Due to the new specification in terms of final qualifications, the point density is not specified, but in practice it will be between about 8 to 20 points per sq m. Apart from the laser point cloud itself, also grids must be delivered with grid sizes of 0.5 and 5 meter (see section 5.2).

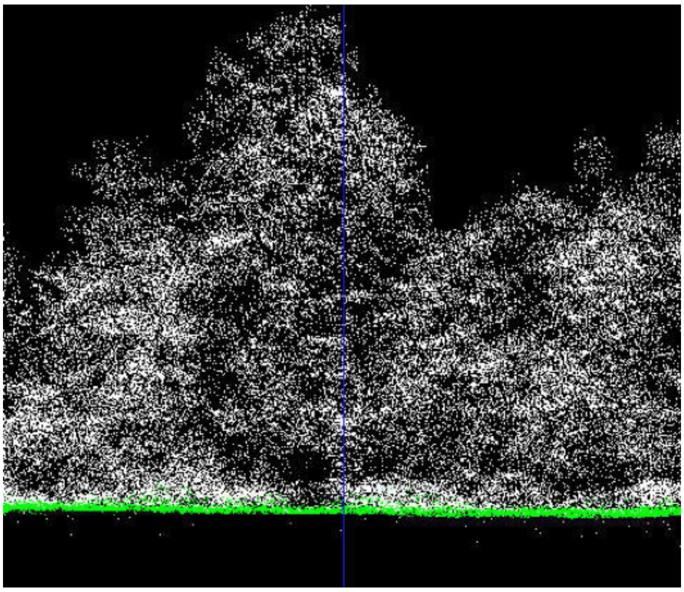


Figure 5. Classification of terrain (green) in the point cloud of an area with dense vegetation. Because of this vegetation, the AHN requires the acquisition to be in the winter (image Fugro Aerial Mapping).

Hydrologic modelling for water management, an other main task of water boards, imposes tight requirements on the accurate description of the terrain surface of the height model. Because of this, the data not representing the real terrain ('maaiveld') must be filtered out and delivered as a separate dataset (cf. Figure 5). Because leafs intercept the laser beam, the AHN requires acquisition of the data in the winter period, from December to April. The criteria for this classification process have been described elaborately and the data set is checked on this to the full extent by the quality control contractor because of its importance.

3.3. Realisation of AHN-2 from 2008 to 2012

After the pilot, done in 2007, the organisation and process were evaluated and it was decided that the next generation AHN would be viable. In 2008 the first part of this new AHN-2 was acquired for seven water boards. In a cycle of five years, the whole of the Netherlands will be covered. In 2012 the data acquisition will be completed, leading to an up-to-date high-precision high-density AHN-2 available for the whole of the Netherlands in 2013 (cf. Figure 6).

4. Assessing the potential of AHN-2 data for dike management

4.1. Workgroup to improve practical application of laser altimetry

In 2005, after the sliding of the peat dike of Wilnis and other calamities, an inventory for the Foundation for Applied Water Research (STOWA) showed that laser altimetry was the most promising technique in acquiring data that could improve dike management fast and effectively. STOWA installed a workgroup to improve the practical application of laser altimetry by water boards. This Workgroup Large-scale Laser altimetry (Werkgroep Grootschalige Laseraltimetrie, WGL) consisted of water board employees experienced in the application of laser altimetry and external experts.

The workgroup investigated the technological developments in laser altimetry and set up requirements to laser altimetry products, based on an analysis of existing tendering documents, a geometrical analysis of the directions for the legal assessment of dikes (VTV) and the experience of dike managers. This indicated that the AHN-2 specifications would indeed suffice to a large extent for dike monitoring.

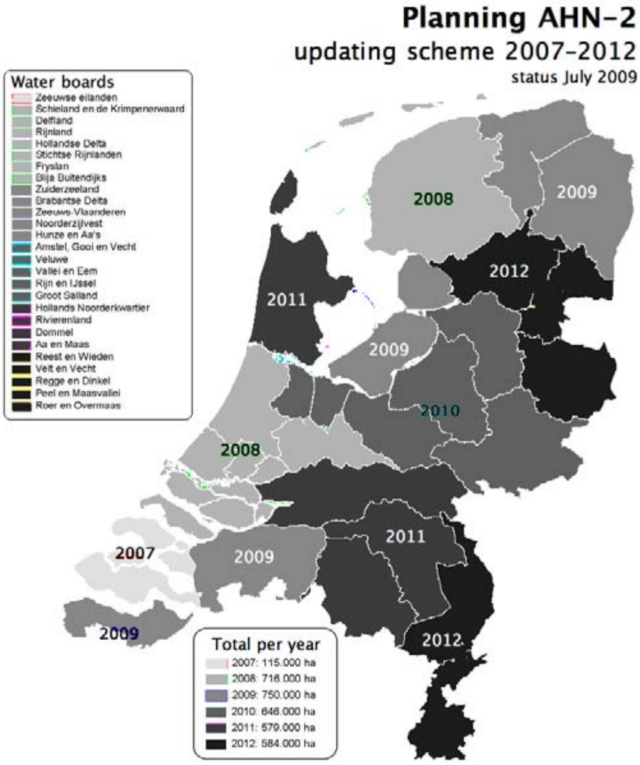


Figure 6. Planning of the next generation AHN (the AHN-2) per water board.

However, dike managers required one extra product: aerial photography of the scene during the laser acquisition. They need this to be able to identify and interpret objects or artefacts that seem to exist in the laser data. Because dike safety can depend on tiny details and is so important for the Netherlands, these detailed requirements must be set. This can pose severe limitations to the laserscanning contractor, because it makes laserscanning by night impossible, when atmospheric and weather conditions are often best. Therefore it was agreed upon that the requirement on aerial photography in AHN-2 will be that aerial photography must be acquired within one week from the laser acquisition. To limit costs, no requirements were set on mosaicing, colour correction and smoothing: the photographs must only be geo-referened.

4.2. Comparison of profiles from AHN and terrestrial data

The AHN steering committee supplied laser data of the pilot with WZE to the WGL, while water board WZE supplied terrestrial profiles of dikes. These data were analyzed by the WGL (Figure 7) (Swart et al., 2007 and Swart et al., 2009). The laser data complies to the requirements mentioned before.

A bilinear interpolation of height data from the laser grid of 50 cm was compared to the terrestrial profiles. Figure 8 illustrates that, although it is known that the terrestrial point measurements are more accurate, they give a coarser description of the topography of the dike. The surveyor makes a decision at which characteristic locations he makes a measurement, but his decision cannot be judged from the laser data only. Although the individual laser measurements have a lower accuracy, they provide a better description of the actual topography of the dike.

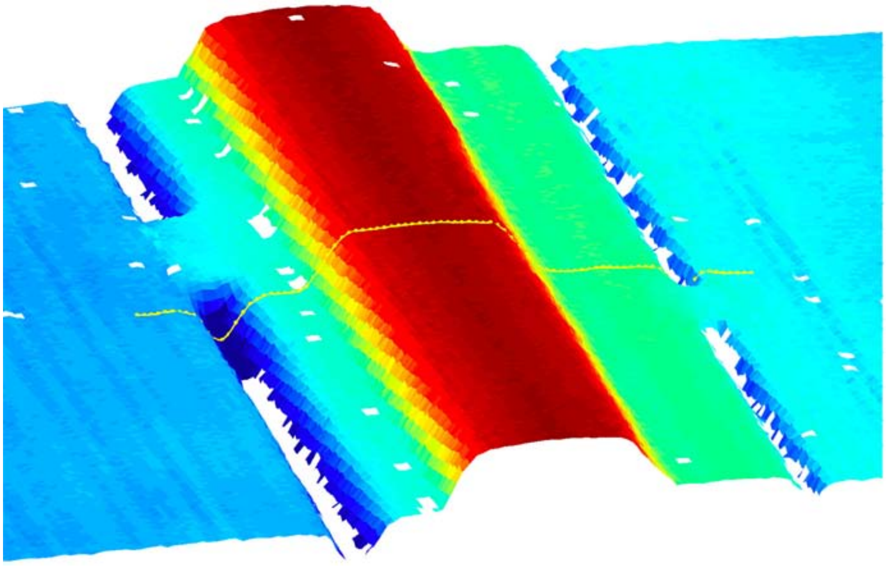


Figure 7. Height model (height exaggerated 2 times) of a 100 meter section of a dike about 2 m high. The 0.5 m laser raster is clearly visible. The yellow dots indicate the location of the terrestrial profile, used for the analysis (data AHN and WZE; analysis Swartvast/STOWA-WGL).

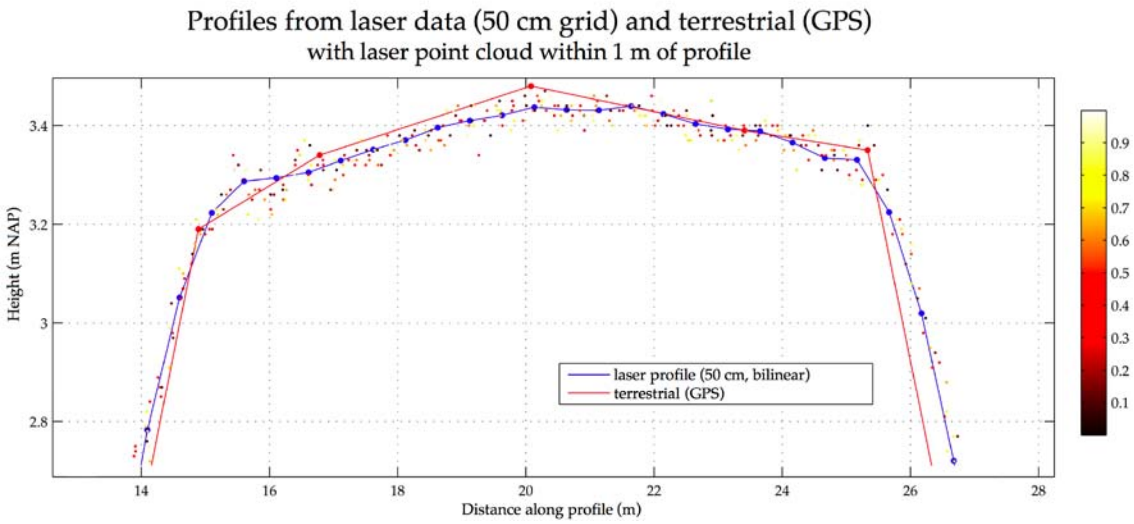


Figure 8. Height profile (exaggerated 6 times) of the top of a dike. The height data, bilinearly calculated from the 50 cm laser grid (blue dots), give a more detailed description of the topography than the terrestrial measurements (red dots). To be able to judge the variations in the individual laser point cloud and their relation with the laser grid, the individual laser measurements are also plotted. The point colour reflects the distance of the point ($0 \leq d \leq 1$ meter) to the profile (see Figure 13), as the legend shows (data AHN and WZE; analysis Swartvast/STOWA-WGL).

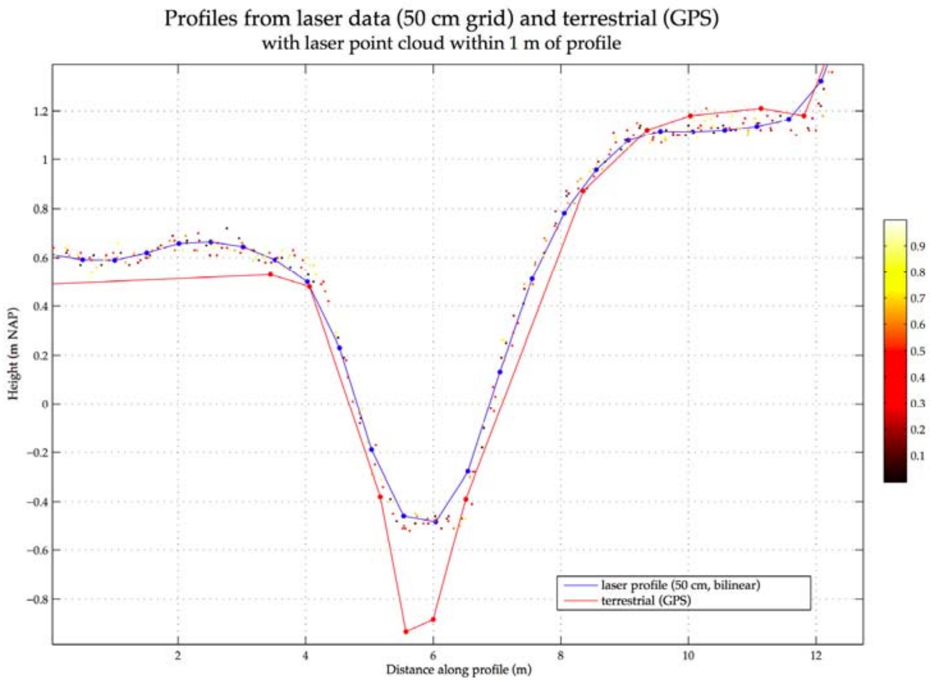


Figure 9. The profile near the ditch at the toe of the dike, plotted as described in Figure 8. The laser data reflect the topography with more detail than the terrestrial profile, although the height makes a significant excursion that cannot be judged without further knowledge. The depth of the ditch can only be measured with a terrestrial technique.

Sometimes significant discrepancies occur between laser and terrestrial data (Figure 9), but it cannot be judged without further knowledge what the reason is. It cannot be stated that in all cases the terrestrial data should be preferred. There is one exception, also illustrated in Figure 9. For the dike stability calculations, the water level in the ditch at the toe of the dike and the depth of the ditch itself is crucial. Both types of measurements cannot be derived from the laser data. The water level can be measured with more accuracy in height as well as location with a terrestrial technique. The ditch bottom cannot be observed with laser at all.

It is considered common sense with water boards that terrestrial data is superior to laser data because of its higher point accuracy, but the practice to use terrestrial profiles to do the quality control of laser data must be advised against. The decision where to do the measurement results in a height variation much larger than the intrinsic accuracy of the terrestrial measurement (this phenomenon is known in geodesy as the 'idealisation accuracy'). The much denser description of the terrain by the laser data makes the lower accuracy acceptable.

4.3. Comparison of stability calculations based on AHN and terrestrial data

For the legally obliged assessment of dikes, also the stability of the dike must be calculated. This is done by calculating the resistance against sliding, based on the best assumption on composition of the dike body, maximum probable water level at both sides of the dike and the profile of its surface. Deltares compared the stability number using both the terrestrial and laser profile. In general, no significant differences were found, except were the ditch geometry was not defined well, as in Figure 9. In Figure 10, the sliding circle with the highest probability is shown. It was concluded that the AHN-derived profiles are suitable for stability calculations.

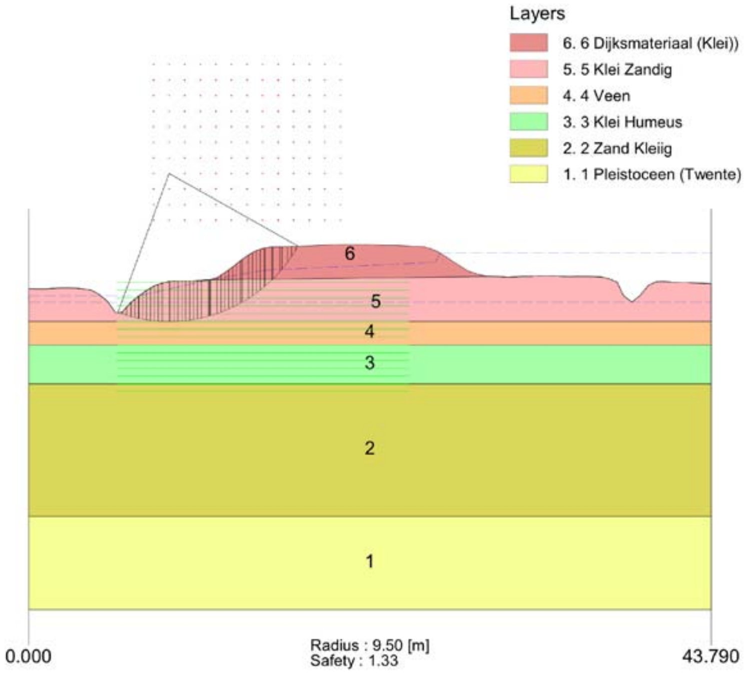


Figure 10. The dike stability calculations reveal the most probable sliding circle, based on the laser profile of the terrain and assumptions on dike body composition and maximum water level. Note that ditch geometry is, in this case, of large importance (Deltares for STOWA).

4.4. Suitability of AHN data for mapping

An important application of laser data for water boards is the establishment of the public register of dikes, waterways and flood defences (legger en beheerregister). Laser height data combined with aerial photography is used to map important lines like middle top line ('middenkruinlijn'), inner and outer toe lines ('binnen- en buitenteenlijnen') and the jurisdictional area of the water board (core zone plus protection zone). Because of the high density of the data and the extension of these lines, the planimetric accuracy of the location of those lines (sometimes bend lines) is much better than the planimetric accuracy of individual laser points. So, also for mapping purposes, the AHN-data suffices.

4.5. Conclusions

It can be concluded that the grid size of 50 cm yields a sufficiently dense data set to reflect the three-dimensional characteristics of dikes to the extent necessary for dike and water management. About 20% of the 17.600 km of dikes in the Netherlands has a height of less than 1 meter and can be almost invisible in the landscape. Only for these dikes, which are of importance to water management despite their small size, the resolution of 50 cm will possibly not suffice; this needs further investigation.

5. AHN-2 products and their usage by water boards

The requirements of water management and dike management lead to a massive dataset (section 5.1). The AHN data products can be divided into three availability levels (section 5.2). Experience and interviews give an interesting picture of the use of AHN products within water boards and, although this picture will not be representative, it will be described in sections 5.3 (grids), 5.4 (point could data) and 5.5 (aerial photography).

5.1. Result of the user requirements: a massive dataset

The requirements of water management and dike management lead to the specification of the second generation of the AHN as described above. Because of the requirements on planimetric object location accuracy and the 50 cm grid size, AHN-2 becomes a massive point data cloud. The aerial photography adds to that. The total number of laser points of the completed AHN-2 point data cloud will be about 400 billion (4×10^{11}); the 50 cm grid will consist of 135 billion grid cells.

5.2. Products and availability levels

The AHN comprises several products, which in combination will serve almost all height model needs of the water boards and Rijkswaterstaat and potentially many other users. The products are based on the user requirements and actual use and usability of the products. For example, because most users consider a grid to be far easier to handle than a point cloud, both point cloud and grid data are produced and offered. Furthermore, because for many applications a 5 meter grid contains sufficient detail and because it is, due to its smaller file size, easier to handle than the 0.5 meter grid files, both a 0.5 and 5 meter grid are offered.

In this section the AHN-2 products are described. The products can be divided into three classes or levels. Each product level has its own availability and pricing.

The lowest level of products is generally available to 'the rest of the world' and can be ordered from the data service desk of Rijkswaterstaat via the ahn.nl website. It can be divided into laser points and grids. The laser point data cloud is separated into two products: points representing the terrain according to the extensive and strict definitions described in the product specification, and all other, non-terrain points like buildings, vegetation and other objects. Both products are delivered as ascii xyz for compatibility reasons. For a typical 130,000 hectare water board, this is 0.4 TB of data in total.

The 'rest of world' grid data consists of four data sets. The first data set contains the terrain-filtered data, resampled to a grid with cell size of 50 cm. The second data set is almost the same, but the occasional gaps caused by an irregular point distribution should be filled in. An example is shown in Figure 11. Even if the average point density complies to the requirements, the irregular spacing of scan lines can lead to grid cells that do not contain a single laser point. In these cases the gap must be interpolated. In all other cases, the AHN does not allow interpolation because this can be regarded as creation of non-genuine height data.

The third 50 cm grid is calculated by resampling all laser points, terrain as well as non-terrain. This data set will, apart from water, cover almost the whole surface and contain buildings, objects and dense vegetation. Finally, the fourth grid is a resampling of the terrain-classified laser points to a grid with a cell size of 5 meter. The grids are delivered as ascii grid. For a typical water board of 130,000 ha, the four grids total to about 100 GB. As of 2010, the grids will be delivered in the ESRI .adf binary format and hence take up less space.

The intermediate availability level of AHN products consists of products only delivered to the owners of the AHN: the water boards and Rijkswaterstaat. This category actually consists of only one product: the aerial photography, intended to identify and interpret objects or artefacts in the laser data. The aerial photography cannot be ordered by other customers and will in general only be delivered to the target water board. For a typical water board, this data set will take up about 0.5 terabyte, depending on the way the contractor delivers and compresses it.

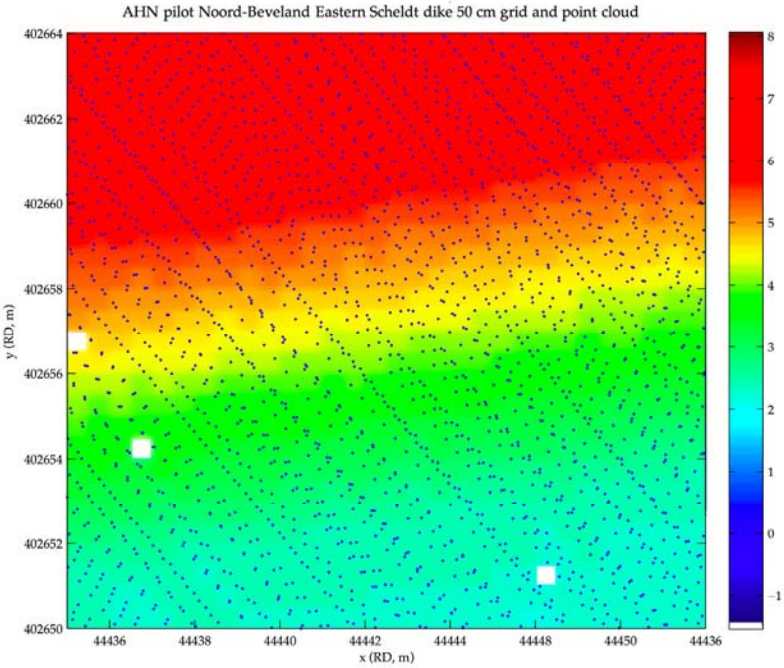


Figure 11. Example of grid cells containing not a single laser point, although the average point density is sufficient. This effect is caused by the irregular point distribution, in this case by irregularly spaced scan lines. This particular example was discovered by the WGL and led to the definition of an additional AHN-product (data AHN, analysis Swartvast/STOWA-WGL).

The most extensive product level comprises of products for quality control of the data acquisition. As acquisition and quality control is performed by separate contractors (although of course the acquiring contractor has its own quality control procedures and reports on these in his own quality report), the quality control products must be explicitly defined. After the data is accepted, the quality control products are archived. They are not distributed and cannot be ordered.

This product level consists of three additional laser data point products, delivered in flight strips. This data is used to check filtering, strip adjustment and accuracy in height and planimetry. The first product is the full laser data set. The second product consists of the terrain-filtered laser points and the third product contains the points filtered out. Also several grid data sets are delivered. For each strip overlap, a 50 cm grid of both laser strips and the difference grid is delivered by the contractor. Point density is checked using a point density grid with 1 meter cell size. The quality control contractor can check the tight filtering requirements on a hill shade grid with 0.5 m grid cell size, delivered by the acquisition contractor, although both contractors in general use additional methods to check the filtering. For a typical water board, all control products take up about 1 terabyte.

5.3. Usage of AHN grids by water boards

For almost all tasks, the AHN grids are used. For hydrologic modelling, the full 0.5 m resolution of the AHN-2 is welcome. The true representation of the terrain, and hence the correct filtering or classification of the terrain to the highest level of detail, is crucial. For hydrologic modelling the grids based on filtered data is postprocessed, because no gaps are allowed and, as water can run under an object, the terrain model should reflect this. The last requirement is reflected in the AHN specifications as precise as is possible for a product covering thousands of square kilometers, but the first requirement is not always met. As a product for general use, the AHN is required not to contain height data where there is no information, to prevent pseudo-information from entering into the height model. Topological issues, like the existence of gaps, are left to the end users.

For water area plans, water boards prefer a lack of detail of the height model. This is not only because the small-scale information is not necessary and takes unnecessary processing time, but also because a high level of detail gives rise to questions of civilians that do not make sense for the goal of a water area plan. In fact, the first generation AHN often suffices.

Polder water level decisions ('peilbesluiten') always give rise to much discussion. For farmers a low level is important because only then they can use heavy machinery to work on the fields, but for the typical Dutch peat meadow-land, a ground water table too low causes oxidation of the peat and hence soil subsidence. Lowering the water table since the Middle Ages is actually the reason that water and dike management is such an important issue in the Netherlands (see for a concise overview in English of water management in past and present days: Arnold et al., 2009). Although the AHN is considered to give a representation of the terrain height as good as possible by extensive processing, water boards apply postprocessing to calculate the best mean soil height. This is done by removing all water ways and even slopes from soil top to ditches and any other non-soil objects that might still be present.

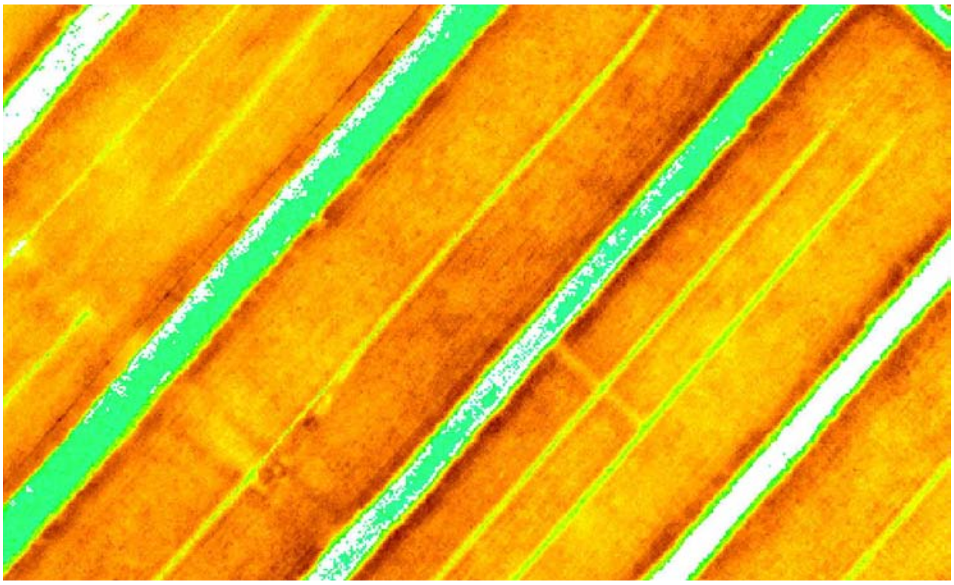


Figure 12. With a suitable legenda, the 50 cm AHN-2 grid shows ditches, drains and other topography of interest to water management with a high level of detail (Hoogheemraadschap De Stichtse Rijnlanden (HDSR)).

For dike management, the AHN data is used to its limits, as shown in section 4 when describing the requirements and assessment by the WGL.

For the public register of dikes, waterways and flood defences, the water boards map several key topography elements, like the middle of the top of the dike ('middenkruinlijn'), and the inner and outer toe lines, using the AHN. This topography is not only used to assess the dike geometry and stability, it also determines the legal zone of jurisdiction of a water board, of importance to managing works that could deteriorate the stability and water resistance of the dike. Sometimes the water boards use the grid to map this topography. As described in the next section, automatic algorithms processing the point cloud data are better suited for mapping. Water boards leave this often to external contractors. An example of mapping dike topography features is shown in Figure 14.

Apart from the applications mentioned before, the AHN is also used for tenders for large-scale maintenance and mowing, for assessing and issuing permits and for enforcing regulations.

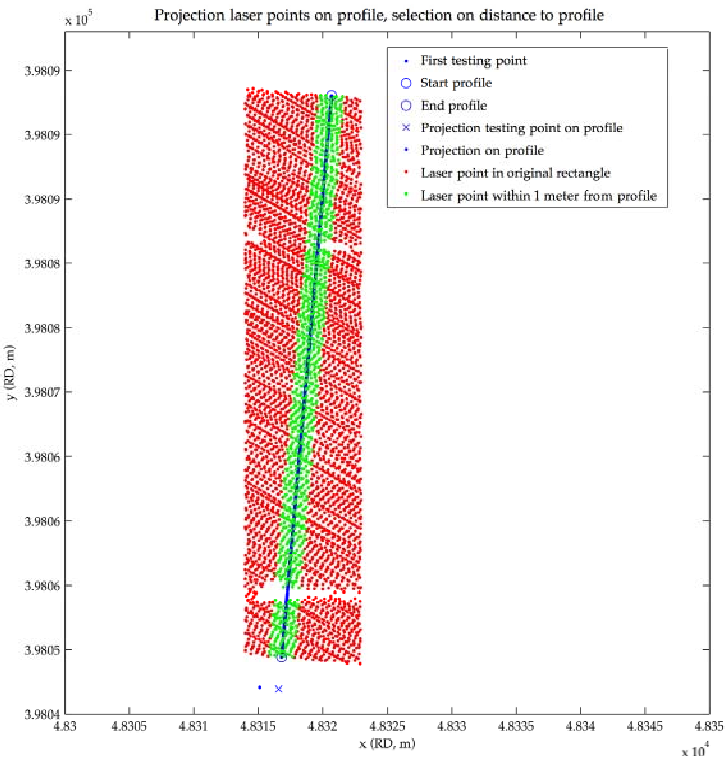
5.4. Usage of AHN point clouds by water boards

Although the requirement for planimetric accuracy of the AHN was deduced analytically from the relation between point planimetric accuracy, point density and point distribution (Vosselman, 2008b), the AHN point cloud products are not used as frequently by water boards as one would expect. Those data is sometimes even kept in the safe deposit and never used.

The main reason for not using of what is actually the original data, is that most data users are far more familiar with grids and the way grids are loaded, presented and analyzed in common GIS software. A second reason is that many standard installations of geo-information systems do not contain modules for handling point cloud data. Another reason is that, in particular because of distribution in ascii format, the sheer size of the point cloud data sets poses an obstacle to users or system administrators to load the data. This however also holds, to a lesser extent, for the grid data. Apart from that, not all storage and networking implementations are suitable to handle these large files and not all workstations and software can handle them in a convenient way.

Furthermore it cannot be denied that also a lack of knowledge, communication and documentation hampers the potential use of point cloud data.

Even the analysis of the suitability of the AHN for dike management by the author for the WGL, described in section 4, was based on the 50 cm AHN grid instead of on the point cloud. This was mainly because the analysis intended to supply support for the discussion whether a 50 cm grid size would suffice. The usage of a grid instead of a point cloud was more or less a pre-



Figuur 13. Although the analysis by WGL was focussed on the grids, the point cloud data was used to be able to judge the variations in the individual laser points compared to the laser grid, as shown by the colour-coded dots in Figures 8 and 9. Laser points were selected in a strip within 1 meter on both sides of the profile.

mise, because without doubt this would be used by water boards. Nevertheless, also the point cloud was used, as shown in Figure 8, 9 and 10 above, but mainly to illustrate the representativity of the 50 cm grid and the excursions the individual points make from the grid. WGL's conclusion that AHN describes dike topography better than terrestrial measurements seems valid even for the grid. The 50 cm cell size of the grid is often detailed enough; it is sometimes even too detailed, as mentioned in section 5.3.

An interesting development in the relation between the finest AHN grid and the point cloud it is calculated from, is that a decade ago the calculation of the AHN-1-grid was sometimes used to improve the height precision of a grid size by averaging several height points, at least when the density increased to several points per standard 5 x 5 m grid cell. For this Inverse Squared Distance Weighting (ISDW) was used (Van Heerd et al., 2000). Nowadays most contractors use the Triangular Irregular Network (TIN) algorithm, that does not average and hence does not improve precision but give a better description of the small-scale topography variations.

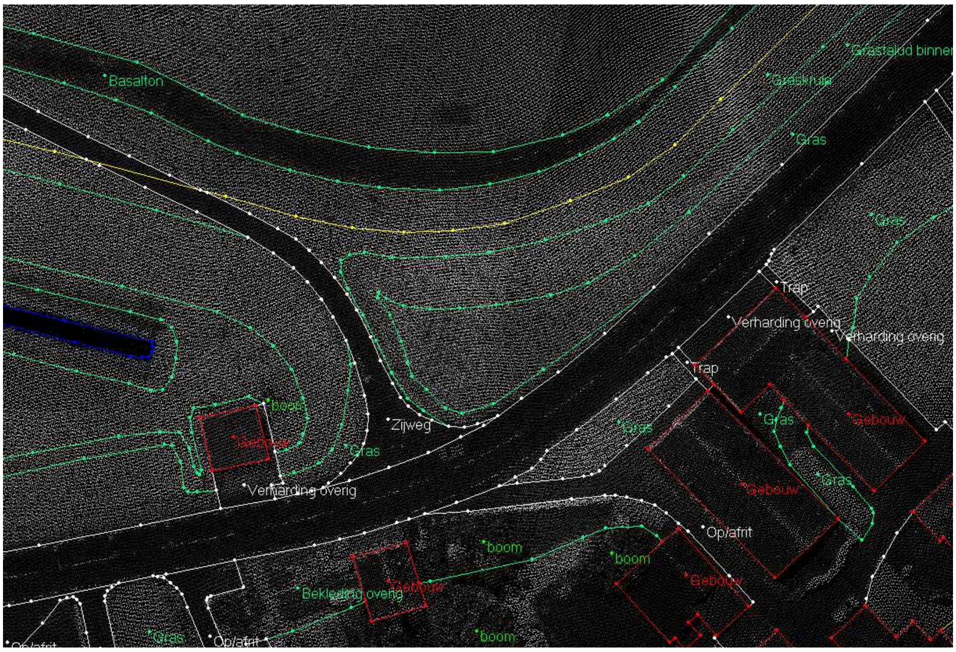


Figure 14. Mapping topography and bend lines using the laser intensity point cloud data and classified height point cloud data (image Fugro Aerial Mapping).

As described in the previous section, an accurate register of key topographic elements for water and dike management is crucial to water boards. For mapping of topography like bend lines it is best to use the point data (see also Brügelmann, 2001). Because of the high density of the data and the extension of these lines compared to the point distance, the planimetric accuracy of the location of those lines is better than the planimetric accuracy of the individual laser points. Because this mapping requires special software, at least if the software automatically determines bend lines from the topography and laser intensity data (as in Figure 14), water boards mainly leave this to specialised engineering agencies.

5.5. Usage of AHN aerial photography by water boards

Although the AHN aerial photography complies to strict requirements which result from the application of the photography as mentioned before, not all water boards use the aerial photography. Some find the photographs difficult to use if they are not mosaiced. Mosaicing and smoothing was not required to limit costs and because it did not seem necessary for dike management. On the other hand, some contractors make mosaics during the quality control process but these are not delivered because they are not required.

Some water boards, together with provinces and municipalities, have a contract for a yearly mosaic. Although this photography is of much higher image quality than the AHN aerial photography and hence is prefered by some water boards, it does not comply to the strict AHN-requirements regarding time of the year and interval between its acquisition and the laser acquisition. A lack of knowledge, communication and documentation seems to lead to fewer use of the aerial photography than would be justified.

6. Conclusions

For large-area, low-detail applications the first generation of the AHN, completed in 2003, can be sufficient. However, the second generation AHN with its upgraded specifications serves the needs of all kinds of applications within the water boards, Rijkswaterstaat and external users. The product requirements were defined in terms of user needs instead of by detailed technical specifications. In 2012 the data acquisition will be completed, leading to an up-to-date high-precision high-density AHN-2 available for the whole of the Netherlands in 2013. The detailed requirements on the exact representation of the terrain leads to products that are suitable for water management, e.g. hydrologic modelling. Meanwhile, an assessment of the pilot data showed that the grid size of 50 cm yields a sufficiently dense and accurate data set to reflect the three-dimensional characteristics of dikes, although maybe not for dikes lower than 1 meter. Hence it was concluded that the AHN-2 can be used to fulfil the legal obligations on the assessment of dike height and stability. The laser measurements provide a better description of the actual topography of the dike than terrestrial profiles do. In this way, the upgraded specifications of the AHN-2 make it possible to unify the acquisition of height data for both water management and dike management.

The AHN product family consists of four grid data sets with a cell size of 0.5 and 5 meter and two point cloud data sets. For the target water board and quality control additional products are available, like aerial imagery. For almost all water management tasks, the AHN grids are used. In general the full 0.5 m resolution is needed. For some applications, this is processed further. The aerial imagery, intended for interpretation of the laser data, is not always used, sometimes because it is not mosaiced.

Most water boards do not use the point cloud data, mainly because most users are far more familiar with working with grids and the software and hardware environment does not always accommodate the convenient use of the point data. But also a lack of knowledge, communication and documentation hampers the potential use of point cloud data. Nevertheless, the 50 cm grid data is in general sufficient to fulfill the user requirements. Only for mapping purposes the point cloud data is used and this is mainly left to external contractors.

References

Arnold, G., N. Kielen, H. Bos, F. van Luijn, R. Doef and R. Goud (2009), Water management in the Netherlands, Rijkswaterstaat Waterdienst.

Berg, P. van den, P.J.L. Blommaart, J.J. Flikweert et al. (2004), De veiligheid van de primaire waterkeringen in Nederland. Voorschrift Toetsen op Veiligheid voor de tweede toetsronde 2001–2006 (VTV), Rijkswaterstaat DWW.

Brand, G.B.M., M.J.E. Crombaghs, S.J. Oude Elberink, R. Brügelmann and E.J. de Min (2003), Precisiebeschrijving AHN 2002, Rijkswaterstaat AGI.

Brügelmann, R., K.I. van Onselen and M.J.E. Crombaghs (2001), AHN basisbestand versus gridbestand. Over de mogelijke meerwaarde van de onregelmatige laserpunten ten opzichte van het 5m-grid, Rijkswaterstaat AGI.

Heerd, R.M. van, E.A.C. Kuijlaars, M.P. Zeeuw and R.J. van 't Zand (2000), Productspecificatie AHN 2000, Rijkswaterstaat AGI.

Swart, L.M.Th., S.J. Flos en W.S. Zomer (2007), Laseraltimetrie voor waterkeringbeheer. Ontwikkelingen, gewenste specificaties, procesbeschrijving en evaluatie AHN-2-proef, STOWA.

Swart, L.M.Th., S.J. Flos, W.S. Zomer and W.G. Epema (2009), 'Meer precisie in laserdata voor waterkeringbeheer', in: Land + Water nr. 4.

Vosselman, G. (2008a), 'Analysis of planimetric accuracy of airborne laserscanning surveys', in: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII, Part B3a, Beijing 2008.

Vosselman, G. (2008b), tender documentation AHN 2010.

* Rens Swart, owner of Swartvast, was contracted by Het Waterschapshuis as project manager of the AHN for 2009.

Recent developments in multi-beam echo-sounder processing – The multi-beam potential for sediment classification and water column sound speed estimation

D.G. Simons, A. Amiri-Simkooei, K. Siemes, M. Snellen
TU Delft, Acoustic Remote Sensing group, the Netherlands

Abstract

The multi-beam echo-sounder (MBES) system allows for unprecedented performance in mapping sea- and river-floors with a 100% coverage. It measures with a single acoustic ping the water depths along a wide swathe perpendicular to the ship track using the travel times of the echo signals received in the acoustical beams. MBES ping rates depend on the water depth, but typically are a few tens of Hz. The MBES opening angle is about 150 degrees and contains several hundred narrow beams, thereby providing high-resolution bathymetric maps. The bathymetry is obtained on-line. Frequently, however, knowledge about the water column sound speeds is insufficient for correctly converting the measured travel times to depths. Here, we present a method for estimating the unknown sound speeds from the MBES data itself. This method can be applied as soon as data along overlapping swathes become available. For this semi-online processing, efficient optimization approaches have been implemented. In addition to the travel times, the MBES also provides measurements of the backscatter strengths in each beam, which are known to contain information about the sediment types. We show results of applying a model-based classification method that employs the MBES backscatter data and discriminates between sediments in the most optimal way. The method has been applied for classification of sediments in a large number of areas. Here, we will show classification results for parts of the river Waal and the North Sea.

1. Introduction

The multi-beam echo-sounder (MBES) system allows for unprecedented performance in mapping the bathymetry of sea- and river-floors. The system measures with a single acoustic ping the water depths along a wide swathe perpendicular to the ship track, thereby covering a large area of the sea- or river-floor at once.

The MBES opening angle is about 150 degrees. Beamsteering at reception allows for determining the (two-way) travel-time of the received signals as a function of angle. Water depths along the swathe can be derived from the combination of travel-time and angle, provided that the local sound speed profile in the water column is known (Beaudoin 2004). Typically, the number of beams amounts to several hundreds, thereby providing high resolution bathymetric maps. It is standard practice to carry out MBES surveys with at least a small overlap between adjacent swathes. The sounding density depends on the MBES opening angle, the sailing speed, the overlap and the water depth. For a water depth of 50 m and a ping rate of 5 Hz, typically 1 – 5 soundings per square meter are obtained. The resulting amount of data acquired per day of surveying is at least a few gigabytes. Still, the bathymetry is determined on-line, and the bathymetric map is established on the fly. However, the increased use of the MBES has given rise to additional requests from the user community, resulting in high demands with respect to e.g. the accessibility of the data. Two of these applications are considered in this paper. The focus, however, lies on the physical principles behind these applications.

The first deals with the possibilities to use the MBES in environments with a highly dynamic water column. To capture the resulting variations in the water column sound speeds, a large number of sound speed measurements are needed. Due to the high costs involved, often only a limited number of measurements are taken. This can result in insufficient knowledge about the sound speeds in the water column, preventing a correct conversion of the measured travel times to depths. The resulting errors in the derived bathymetry may be such that the survey needs to be repeated. An alternative approach is to exploit the redundancy in the MBES measurements, resulting from the overlap between adjacent swathes. The maximum time between measuring two overlapping swathes typically amounts to several hours. Since bottom features, such as mega ripples and sand waves, are not expected to vary significantly on this time-scale, the bottom can be assumed to be stable over the course of the survey. Consequently, the depths as determined from the measured travel times along two overlapping swathes should be the same at equal points on the seafloor. The sound speeds are then estimated by minimizing the difference between the water depths at the overlapping parts of the swathes. This method not only provides the bathymetry for which errors due to insufficient knowledge about the water column sound speeds have been diminished, but it also estimates the water column sound speeds. In principle, the method allows for MBES bathymetric measurements where no sound speed information is acquired as long as the overlap between adjacent swathes is sufficient. This method can be applied as soon as data along overlapping swathes become available. For this semi-online processing, efficient minimization approaches have been implemented.

The second demand from the user community considers the use of acoustic measurements for classifying the sediments. To extract, in addition to bathymetry, also the sediment composition from the MBES measurements would allow for huge cost savings, since it eliminates the need for costly sampling campaigns. Knowledge about the sediment composition is of importance in e.g. marine habitat studies, morphodynamic or sediment transport studies.

It is well-known that the acoustic signals as received by the MBES are affected by the interaction with the sediment. The interaction is dependent on the sediment type. Fine muddy sediments, for example, are known to result in low backscattering, whereas coarse sediments such as gravel result in high backscatter (APL 1994). By employing this knowledge, in theory, the MBES can be used also for sediment classification purposes. However, the process of extracting sediment type from the backscatter measurements requires dedicated processing steps. We have developed a model-based method that employs the MBES backscatter measurements for sediment classification. The method fully accounts for statistical fluctuations in the backscatter intensity and consequently discriminates between sediments in the most optimal way. It estimates both the number of seafloor types present in the survey area and the probability density function for the backscatter strength at a certain angle for each of the seafloor types. Other MBES classification approaches have been developed both by commercial companies and universities. (Simons 2008) (Clarke 1994), (Hellequin 2005) (Canepa 2005) All classification approaches currently can only be applied in a post-processing step and are not yet automated.

Section 2 of this paper describes the approach taken towards eliminating the sound speed induced errors in the bathymetry by estimating the prevailing water column sound speeds. The performance of the method is demonstrated by applying it to MBES data that have been acquired in the Maasgeul, the Netherlands, which is known to be highly variable with regards to the water column sound speed. Section 3 presents the model-based classification method based on the MBES backscatter data. The method has been applied for classification of sediments in a large number of areas. Here, we will present classification results for parts of the river Waal and the North Sea.

2. Compensation of multi-beam echo-sounder (MBES) bathymetric measurements for errors due to the unknown water column sound speed

2.1 Description of the approach

MBES systems emit acoustic pulses in an opening angle of 1 to 2 degrees in along-track direction, and about 150 degrees in the across-track direction. Beamforming in across-track direction is applied to determine the corresponding two-way travel-times for a selected number of arrival angles. Water depths along the swathe, spanned by the across-track opening angle, are determined from the combination of two-way travel-time and angle. Hereto, either propagation along straight sound rays is assumed, or in case the curvature of the sound rays cannot be neglected, ray-trace calculations are carried out.

Inaccurate knowledge about the water column sound speeds results in an erroneous bathymetry in two ways:

1. Errors in the beamsteering process. In case the actual sound speed deviates from the measured sound speed, the actual beamsteering angles differ from the beamsteering angles aimed for, and are unknown.
2. Errors in the conversion from the angle and travel-time combinations to water depths along the swathe.

Figure 1 shows the geometry of a typical MBES survey, consisting of a series of tracks sailed parallel to each other. Track distances are such that a certain overlap exists between adjacent swathes.

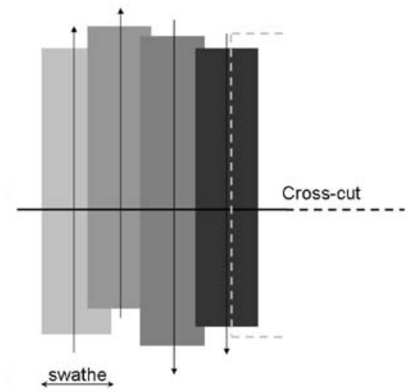


Figure 1. Schematic of an MBES survey. Arrows indicate sailing directions, grey-shaded rectangles the area measured per track.

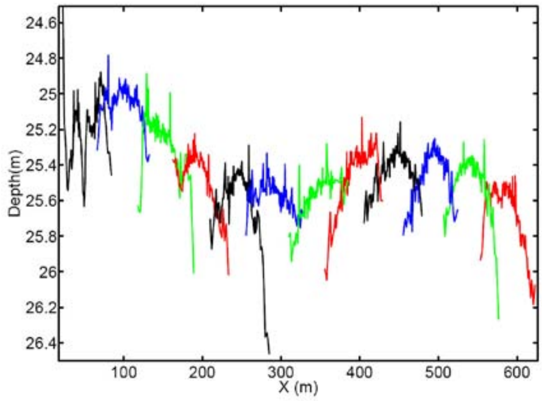


Figure 2. Example of 'droopy' effects along a cross-cut. The MBES measurements were carried out near the entrance to the harbor of Rotterdam.

Figure 2 shows an example of the typical bathymetric behavior along a cross-cut, such as indicated in Figure 1, in case erroneous sound speeds are used. The area in which these MBES were taken is located close to the entrance of the Rotterdam harbor, where mixing of fresh and salt water occurs. The number of parallel tracks amounted to 12. The bathymetry was estimated from the measured travel-times, employing all sound speed information available, i.e., sound speeds measured at the transducer head for the beamsteering and a single sound speed profile for calculating the sound propagation through the water column. The colors indicate the bathymetry as estimated for each of the tracks. Differences in water depths along the overlapping parts of adjacent swathes amount to almost 0.5 m.

Currently, efforts to reduce these effects mainly aim at collecting additional sound speed profile measurements. In addition, research has been carried out to assess the potential of using oceanographic models for predicting the prevailing sound speeds. (Calder 2004) In this paper, a different approach is taken, where the sound speeds are estimated from the MBES data and no additional sound speed profile measurements are needed.

The method that has been developed for eliminating these effects fully exploits the redundancy of measurements in the overlap region between two adjacent swathes. The measurements are the two-way travel-times per beam and swathe. For estimating the unknown bathymetry and sound speeds, a model is required that predicts these two way travel times, given a set of values for the unknowns. Hereto, the seafloor is modeled with an interpolated grid function, with the water depths at the grid positions being unknowns that need to be determined. We aim to minimize a function that represents the mismatch between measurements and model output, i.e.,

$$E = \sum_{k=1}^{S} \sum_{j=1}^{N} \left(t_{k,j} - T_{k,j} \right)^2 \tag{1}$$

where N and S are the total numbers of MBES beams and swathes, respectively. The measured two-way travel-times are denoted by $T_{k,j}$. The modeled two-way travel-times are denoted by $t_{k,j}$. The model that calculates $t_{k,j}$, accounts for both the effect of sound speed on the beamsteering and on the propagation through the water column. The unknowns are the sound speed profiles for each of the swathes and the bathymetry. These unknowns should be estimated such that E becomes minimal.

For the minimization we have considered two different optimization techniques:
– The method of Differential Evolution (DE) (Snellen 2008);
– Gauss-Newton (GN).

DE is a global optimization technique and can be seen as a modified version of a Genetic Algorithm. It does not make use of derivatives, but searches within all possible solutions for promising solutions. DE has been applied extensively, showing good to very good performance in locating the global optimum of a function. The advantages lie in its robustness (i.e., its insensitivity to the shape of the energy landscape) and the fact that it does not pose requirements on the behavior of the function considered. The drawback is that it requires a significant number (1000 – 3000) of forward model calculations, in this case model calculations for $t_{k,j}$.

The advantage of GN lies in its efficiency, i.e., it requires much less computations. The drawbacks are that it is a local method, likely to end up in a local minimum in case multiple minima exist, and that it requires expressions for the derivatives of the function with respect to the unknown parameters (unless calculated numerically). These derivatives are dependent on the assumed model for $t_{k,j}$. For the situation at hand, for example, each parameterization of the water column sound speed results in different expressions for $t_{k,j}$.

Based on a series of DE inversions for different parameterizations of a shallow water sound speed profile (constant sound speed, linearly varying sound speed, and a water column consisting of two different layers, each with a different sound speed), all were found to give similar results for the estimated bathymetry. Therefore, for GN we have assumed a constant sound speed throughout the water column. This results in one unknown sound speed for each of the tracks. The unknowns (~250) to be optimized are thus the depths at the grid positions and the sound speeds for each of the swathes.

The optimization procedure is as follows: For the seafloor, we use a fixed grid of closely spaced horizontal positions, denoted X_n in the across-track direction. At every position X_n, Z_n denotes the corresponding water depth. These Z_n are part of the unknowns to be estimated. Between the grid points the depth is interpolated linearly. For every beam j at angle $\theta_{k,j}$, the point where the acoustic beam impinges on the model seafloor is denoted as ($x_{k,j}$, $z_{k,j}$). Figure 3 shows a schematic overview of this model.

Let the MBES be located at ($X_{k,MBES}$, $Z_{k,MBES}$), then the function for $t_{k,j}$ can be derived by calculating the intersection between the sound ray and the line between the grid points, as illustrated in Figure 3.

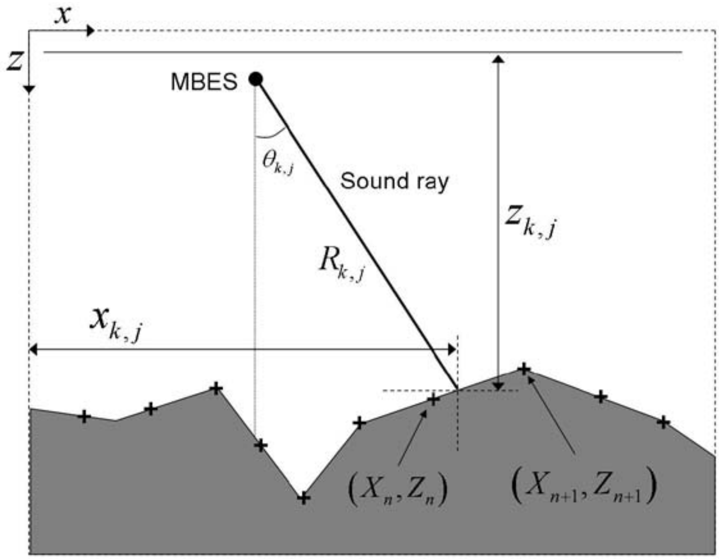


Figure 3. Illustration of the model for $t_{k,j}$. On the right side of the MBES we show the path of a sound ray through water column. The ray impinges at the seafloor in between the grid points X_n and X_{n+1} at coordinates ($x_{k,j}$, $z_{k,j}$). $t_{k,j}$ is calculated as $2R_{k,j}/c$.

The position at which these two lines intersect is given by:

$$x_{k,j} = \frac{\dfrac{X_{k,MBES}}{\tan\theta_{k,j}} + Z(X_n) - Z_{k,MBES} - X_n \dfrac{Z(X_{n+1}) - Z(X_n)}{X_{n+1} - X_n}}{\dfrac{1}{\tan\theta_{k,j}} - \dfrac{Z(X_{n+1}) - Z(X_n)}{X_{n+1} - X_n}} \tag{2}$$

and

$$z_{k,j} = \frac{x_{k,j} - X_{k,MBES}}{\tan\theta_{k,j}} + Z_{k,MBES} \tag{3}$$

Values for $t_{k,j}$ are calculated employing the water column sound speed c_k as

$$t_{k,j} = \frac{2\left(x_{k,j} - X_{k,MBES}\right)}{c_k \sin\theta_{k,j}} \tag{4}$$

where c_k is the sound speed in swathe k.

For each realization of the unknowns, i.e., the sound speeds c_k and the depths Z_n, values for $t_{k,j}$ are calculated. Minimizing E of Eq. (1) implies a search for those values of the unknowns that minimize the difference between the measured ($T_{k,j}$) and calculated travel times ($t_{k,j}$). To solve for Eq. (1) in a least-squares sense requires the iterative Gauss-Newton approach. For this, the expressions for the derivatives of $t_{k,j}$ to all unknowns have been determined. In general, two to three iterations were found to be sufficient for localizing the optimal values for Z_n and c_k. Either the optimized depths can be used directly as the corrected bathymetry, or the optimized sound speeds can be used to recalculate the bathymetry for each beam and each ping.

In (Snellen, 2009) simulations are presented to assess the required overlap. It was found that both an increasing noise level and a decreasing overlap result in increased deviations of the estimations from the true bathymetry and sound speeds, but that still the deviation is limited, with mean errors in the sound speed and bathymetry seldom exceeding 1 m/s and 0.2 m.

2.2 Results

The method has been applied to MBES data that were collected in the Maasgeul, the Netherlands. The black rectangle in Figure 4 indicates the area of the measurements. The data have been collected during a standard survey, during which a single sound speed profile was measured. The sound speed values at the MBES transducer were measured continuously.

The left plot of Figure 5 shows the bathymetry as determined from the MBES measurements, where all available sound speed information was used for converting the measured travel times to water depths. Artifacts running parallel to the sailing direction are clearly visible. These show a typical 'droopy' structure, indicating that they are caused by errors in the conversion from travel times to water depths due to insufficient knowledge about the prevailing sound speed profiles. These droopy effects are most pronounced in the outer beams. However, the bathymetric errors are also present directly underneath the sonar.

The plot on the right of Figure 5 shows the bathymetry as obtained after application of the refraction-correction method described above. Clearly the artifacts have been eliminated. The beam trawl marks (depths of several dm's) are still visible.

It can be concluded that the proposed method strongly increases the potential of the MBES to accurately measure the bathymetry in environments with a strongly varying water column. The availability of efficient optimization techniques allows for application of the bathymetry correction in real time, i.e., as soon as the first overlaps become available.

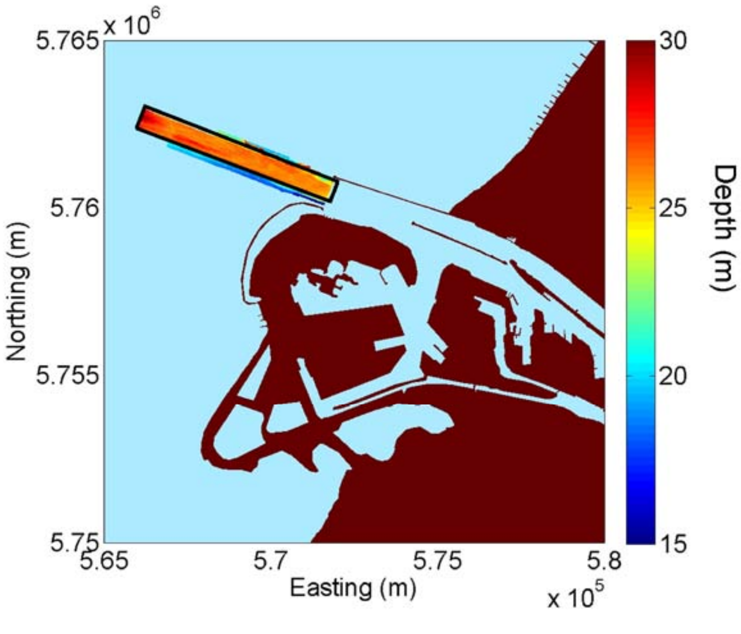


Figure 4. The Maasgeul area (close to the Rotterdam harbor) where the MBES measurements were taken.

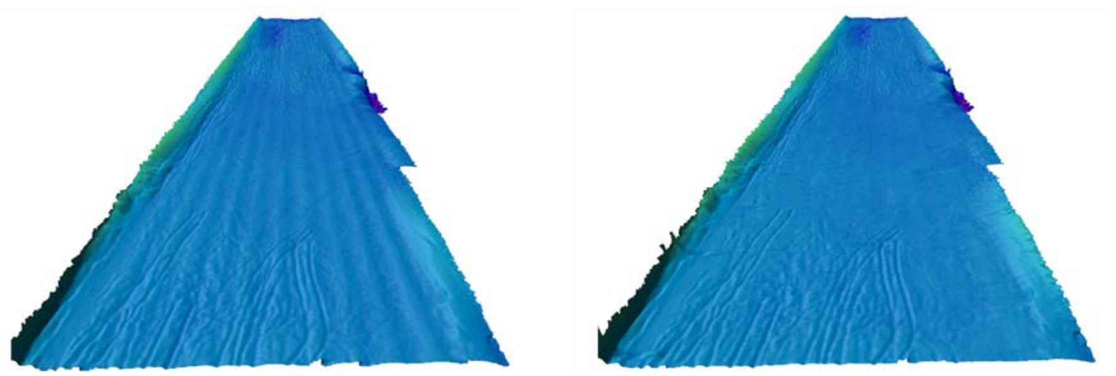


Figure 5. Left plot: Bathymetry as determined from the MBES measurements before application of the proposed method. Right plot: bathymetry obtained from application of the correction method.

3. Sediment classification with the multi-beam echo-sounder

3.1 Description of the method

The approach presented classifies the sediments in an area based on the measured backscatter. The backscatter intensity varies with incidence angle. This angular dependence masks effects of variation in sediment type and morphology in the backscatter images. Therefore, MBES systems apply corrections to the measured backscatter intensities to eliminate this angle dependence, e.g. by assuming Lambert's law. The backscatter images obtained from the MBES after angle correction are comparable to those obtained with a side-scan sonar system (SSS). These maps can be used for classification purposes by resolving textures or spatial variability in the data., e.g. as in (Blondel 2009). On the other hand, the variation of the backscatter strength versus angle with sediment type can potentially be used for classification (Simons 2008), (Clarke 1994) (Canepa 2005). A problem in this latter approach arises for areas where the sediment type varies along the swathe. In this case, it is difficult to discriminate between the angular variation itself and real sediment type variation along the swathe.

Therefore, we have selected an alternative approach. The approach proposed here employs the backscatter data still containing the angular dependence. It is, therefore, not dependent on the corrections applied to eliminate this angular dependence. However, instead of using the angular behavior of the backscatter strength as the classifying parameter, it uses backscatter measurements per angle. Additionally, it accounts for the ping-to-ping variability of the backscatter intensity that partly masks the information about sediment characteristics. The classification method employs the backscatter values (in dB) per receiver beam, i.e., backscatter values that have been obtained from averaging (or filtering) over the N_s independent scatter pixels in a receiver beam. Corrections for propagation losses and footprint are applied, and backscatter values are provided for each of the MBES beams.

Due to the small pulse length T_p employed by the shallow water MBES systems (typically ~100 μs) the signal footprint is also small. The beam footprint is determined by the water depth H, the beam angle and the MBES transducer characteristics, and typically is much larger than the signal footprint. The number of scatter pixels N_s per beam footprint for a beam at angle θ with the vertical is given by

$$N_s(\theta) = \frac{\frac{H\Omega_{rx}}{\cos^2\theta}}{\frac{cT_p}{2\sin\theta}} \qquad (5)$$

with c the water column sound speed and Ω_{rx} the beam opening angle in the across-track direction. This expression only holds for beams away from normal incidence.

If N_s is sufficiently large for the central limit theorem to hold, the backscatter values per beam are distributed according to a normal distribution. For water depths of a few meters, as encountered e.g. in river environments, N_s is not sufficiently large for the assumption of a normal distribution to hold, and the measured backscatter strengths are averaged over a number of subsequent pings and beams to increase N_s (Amiri-Simkooei 2009).

Based on the assumption of normally distributed backscatter strengths per sediment type, we apply the following approach towards classification of the sediments.

Step 1: Nonlinear curve fitting. The algorithm starts by fitting a model to the histogram of measured backscatter strengths. The data consist of all averaged backscatter data as measured for a certain angle (or set of angles in shallow water situations). The model that we fit to the histograms, therefore, consists of a sum of m Gaussian probability density functions (PDFs), each PDF representing a sediment type with mean backscatter strength $\bar{y}_k$ and standard deviation σ_{yk} (both in dB), i.e.,

$$f(y_j \mid x) = \sum_{k=1}^{m} c_k \exp\left(-\frac{(y_j - \bar{y}_k)^2}{2\sigma_{yk}^2}\right) \tag{6}$$

with $f(y_j \mid x)$ the value of the model at backscatter value y_j, and vector x containing the unknown parameters, i.e., $x = (\bar{y}_1, \ldots, \bar{y}_m, \sigma_{y_1}, \ldots, \sigma_{y_m}, c_1, \ldots, c_m)^T$. As a measure for the match between model and measurements we consider the χ^2 value

$$\chi^2 = \sum_{j=1}^{M} \frac{\left(n_j - f(y_j \mid x)\right)^2}{\sigma_j^2} \tag{7}$$

with M the number of bins in the histogram and n_j the number of y_j occurrences. Assuming that the n_j are Poisson distributed, the variances σ_j^2 are equal to n_j. The unknown parameters are determined by minimizing Eq. (7).

Parameter m, i.e., the number of detectable sediment types present, is estimated by carrying out the fitting procedure for an increasing number of m, until further increasing m, no longer results in an improvement of the goodness of the fit. Figure 6 presents an illustration of this procedure. The thick solid black line in the figures corresponds to the histogram of the measures backscatter strengths. The fit is represented by the red line, with the individual PDFs represented by the thin black lines (4, 5, 6, and 7 for the different subplots, respectively). The goodness-of-fit criterion is defined as the reduced χ^2 – value $\chi_\nu^2 = \chi^2 / \nu$ being close to one, where $\nu = M - 3m$ are the degrees of freedom.

Step 2: Acoustic classes identification. After step 1, both the PDFs for each of the sediment types and the number of sediment types are known. We assume all hypotheses to be equally likely. Then, by applying the Bayes decision rule for multiple (m) hypotheses H_i, a sediment type is assigned to each measurement as follows:

$$\text{accept } H_k \text{ if } \quad \underbrace{\max}_{H_i} f(y_j \mid H_i) = f(y_j \mid H_k) \quad \text{with } i = 1, \ldots, m \tag{9}$$

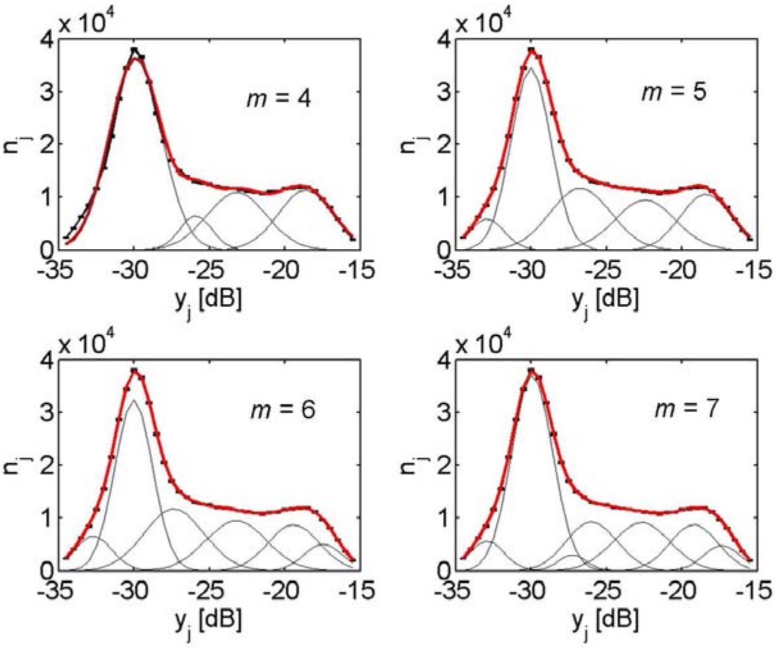


Figure 6. Histograms of measured backscatter data (60 degree beam) fitted to the model for m = 4, 5, 6 and 7, respectively.

This means that we choose the hypothesis that, given the observation *y*, maximizes the likelihood $f(y\,|\,H)$. We, therefore, have to determine the intersections of the *m* normal PDFs, resulting from the fitting procedure of Step 1. This results in *m* non-overlapping acceptance regions A_k.

Step 3: Assigning sediment type to acoustic classes. Now, a sediment type needs to be assigned to each of the acceptance regions. The result of this step can be accomplished by a comparison of the $\bar{y}_k$ values with a combination of data found in the literature, model outcomes and knowledge of the surveyed area based on e.g. cores or samples.

Step 4: Quality assessment. Based on the so-called decision matrix of the multiple-hypothesis-testing problem, probabilities of incorrect decision are determined.

Step 5: Mapping. By plotting sediment type versus position, e.g. with different colors representing different sediment types, a classification map of the area is obtained.

3.2 Results

The approach described in section 3.1 has been applied to MBES data acquired in a large number of different areas, with different sediment types. For the current paper, we present results for two different areas.

The first area is the Cleaver Bank area in the North Sea, the Netherlands. The upper plot of Figure 7 shows the bathymetry in the area, indicating water depths from ~30 to ~60 m. The data were acquired in 2004. The MBES used for the measurements was an EM3000 dual-head system, working at a frequency of 300 kHz. The pulse length amounts to 150 μs. The opening angles in both the along-track and across-track direction amounts to 1.5°. In total 254 beams are formed by the system. In practice, however, the number of beams at which measurements are taken is less and typically amounts to 160.

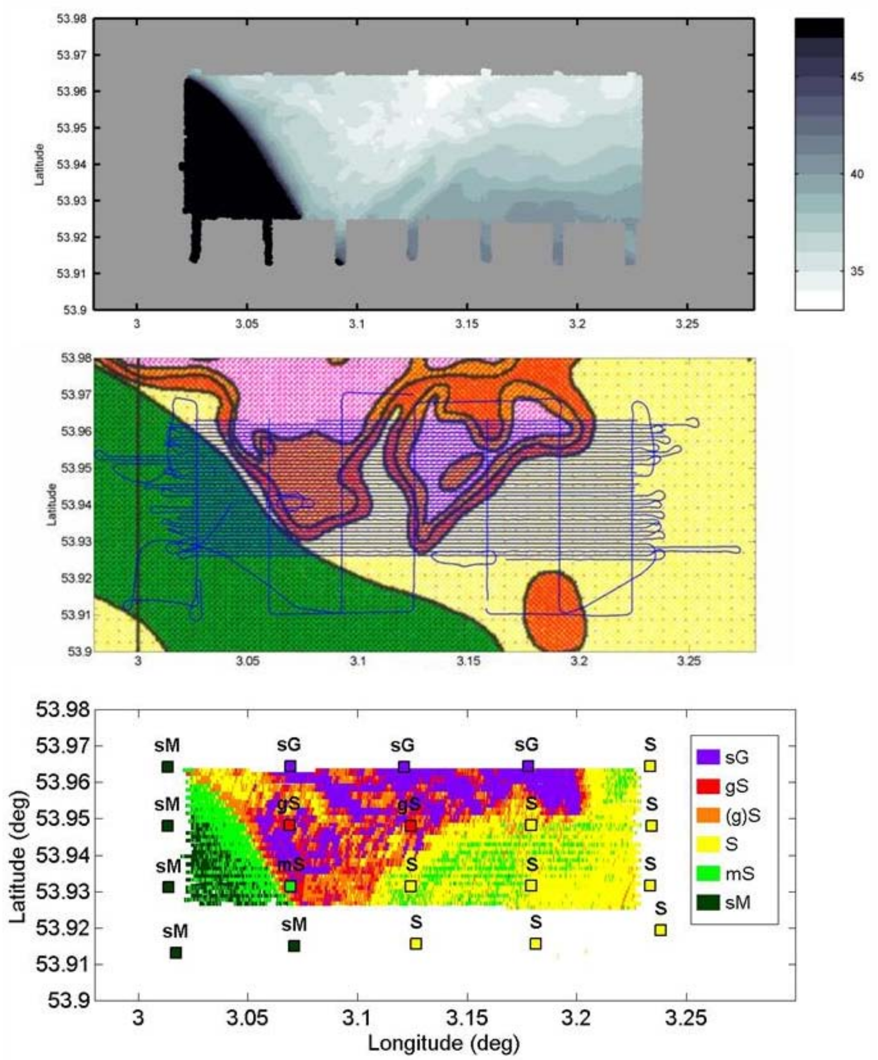


Figure 7. Upper plot) The bathymetry in the Cleaver Bank area. Middle plot) The tracks sailed during the survey, superimposed on a geological map of the area. Lower plot) Classification of the sediments by application of the proposed classification method. Also indicated is the Folk class as a function of position as determined from the grabs. The sediment types present are, i.e., sandy gravel (sG), gravelly sand (gS), slightly gravely sand ((g)S), sand (S), muddy sand (mS), and sandy mud (sM).

The data were provided as raw data files (4.5 GB) from which the backscatter strength as a function of angle and position has been extracted. The middle plot of Figure 7 shows the tracks sailed during the survey. The ping rate is about 5 Hz. The survey lasted for about 32 hours, resulting in ~600.000 pings collected during the survey. For the water depths in the area which range from 30 to 60 m, this corresponds to a 1 – 5 measurements per square meter.

For each ping the backscatter strength at a predefined angle is selected. By applying the method as described in the previous section it is found that, based on the backscatter strengths, 6 sediment types could be discerned in the surveyed area. Samples of the sediment were employed to assign a sediment type to each of the acoustic classes. The resulting classification map is shown at the bottom of Figure 7. Also shown in this plot are the sediment types according to the samples. Clearly, the acoustic classification method reveals a distribution of the sediments over the area that is in good agreement with the distribution indicated by analysis of the sediment samples.

Figure 8 shows the classification map for the second MBES data set, acquired along a part of the river Waal, close to St. Andries. Again the system is a 300 kHz MBES with a maximum of 254 beams. The ping rate is about 35-40 Hz, resulting in ~2 million pings in the surveyed area. Sediments in this area are coarser then those at the Cleaver Bank. As with the Cleaver Bank area, the acoustic classification results are in very good agreement with the sediment samples.

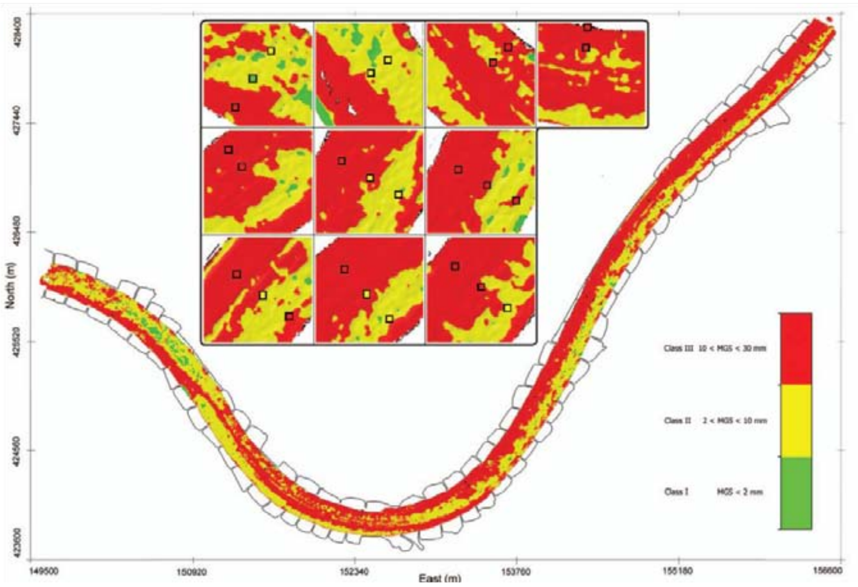


Figure 8. Acoustic classification map for an area of the Waal, close to St. Andries. The three colors indicate the three acoustic classes. The results of the sample analysis are shown by the squares. MGS indicates the mean grain sizes.

4. Summary and conclusions

In this paper two methods are described that focus on improving the performance of MBES systems. For both, efficient access to the individual data points within the large number of measurements is essential.

The first aims at eliminating errors in the bathymetry due to erroneous sound speed information by employing the overlap between adjacent MBES swathes. In principle, this method allows for MBES surveys where no information regarding the prevailing sound speeds is acquired. The approach has been optimized with regards to computational efficiency. Consequently, the method can be employed in real-time during the survey. The only requirement is that there exists overlap between adjacent swathes.

In addition, the performance of the MBES with regards to classification of the sediments is demonstrated. For the research presented in the paper, the classification has been carried out after the survey. However, for practical situations where it might be of interest to establish a picture of the sediment distribution while surveying, another approach can be selected, where each new measurement is added to the previous measurements. The fitting procedure has then to be repeated during the survey, thereby building up and refining the overview of the sediment distribution in the area of the survey.

The combination of high-resolution bathymetry and high-resolution classification results can provide important new insights in the mechanisms governing the sediment distribution. Further future developments in this field will be to fully exploit the entire time series, i.e. the complete received signal per beam, for improved classification. This requires storing the received signal for each ping and each beam, whereas currently the signal is reduced to two parameters, being the two-way travel time and backscatter strength. The advantage of obtaining more information is, therefore, counteracted by a significant increase in data rate with at least a factor of 100 to 1000. Allowing classification approaches that employ the complete signal per beam to be applied real-time is expected to require a significant research effort.

Acknowledgements

We would like to thank Ben Dierikx, Simon Bicknese, and the Dutch Directorate-General for Public Works and Water Management for fruitful discussions and valuable MBES data sets. Also we would like to thank the Dutch Hydrographic Service for providing us with MBES data. This work is financially supported by the Dutch Directorate-General for Public Works and Water Management, the Netherlands Geodetic Commission, and the Technical University of Delft.

References

APL-UW High-frequency Ocean Environmental Acoustic Models Handbook, Technical Report, APL-UW TR 9407, AEAS 9501, October 1994.

Amiri-Simkooei, A.R., Snellen, M., Simons, D.G., ‘Riverbed sediment classification using MBES backscatter data’, Journal of the Acoustical Society of America 126, 2009, p. 1724-1738.

Beaudoin, J., Hughes Clarke, J.E. and Bartlett, J.E., ‘Application of surface sound speed measurements in postprocessing for multi-sector multibeam echosounders’, International Hydrographic Review 5, 2004, p. 26-31.

Blondel, Ph., and Gómez Sichi, O., ‘Textural analyses of multibeam sonar imagery from Stanton Banks, Northern Ireland continental shelf’, Applied Acoustics 70, 2009, p. 1288–1297.

Calder, B., Kraft, B., de Moustier, C., Lewis, J., and P. Stein, ‘Model-based refraction correction in intermediate depth multibeam echosounder survey’, In Proceedings of the Seventh European Conference on Underwater Acoustics, Delft, the Netherlands, 2004, p. 795-800.

Clarke, J.H., ‘Toward remote seafloor classification using the angular response of acoustic backscattering: a case study from multiple overlapping GLORIA data’, IEEE J Ocean Eng 19, 1994, p. 112-26.

Hellequin, L., Boucher, J., Lurton, X., ‘Processing of high-frequency multibeam echo sounder data for seafloor characterization’, IEEE J Ocean Eng 28, 2005, p. 78–89.

Canepa, G., and Pouliquen, E., ‘Inversion of geo-acoustic properties from high frequency multibeam data’, in Proceedings of boundary influences in high frequency shallow water acoustics, Bath, UK, 2005, p. 233-40.

Simons, D.G. and Snellen, M., ‘A comparison between modeled and measured high frequency bottom backscattering’, in Proceedings of the European Conference on Underwater Acoustics, Paris, France, 2008, p. 639-644.

Snellen, M., and Simons, D.G., ‘An assessment of the performance of global optimisation methods for geo-acoustic inversion’, Journal of Computational Acoustics 16, 2008, p.199-223.

Snellen M., Siemes K., Simons, D.G., ‘An efficient method for reducing the sound speed induced errors in multi beam echo sounder bathymetric measurements’, in Proceedings of the Underwater Acoustic Measurements Conference, Nafplion, Greece, 2009, p. 1155-1162.

Storage and analysis of massive TINs in a DBMS

H. Ledoux
TU Delft, GIS technology group, the Netherlands

Abstract

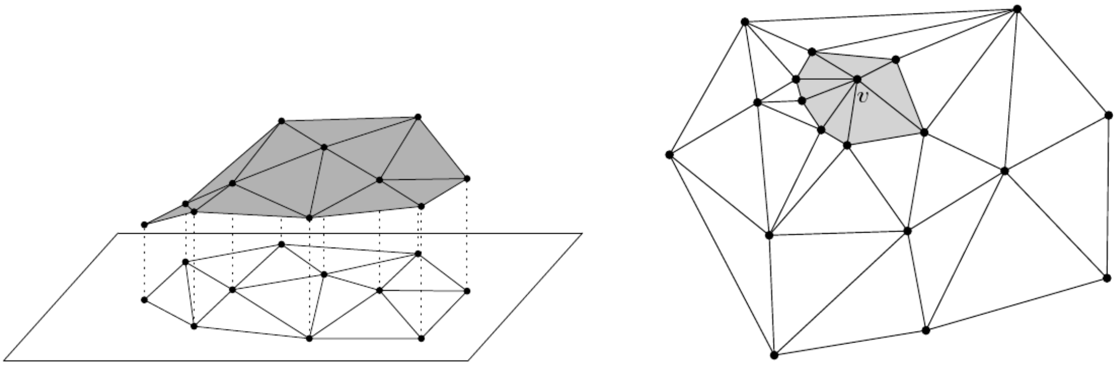
Advances in technologies to collect elevation data are far superior to the advances to store and process the resulting point clouds. This paper investigates the use of a DBMS to store and manage not only the points coming from massive point-based datasets, but also a TIN of these points. TINs are used as a support structure to implement processing and manipulation operators. I discuss in the paper why storing efficiently a TIN in a DBMS is a complex task, and I propose a new solution. It does not store explicitly triangles, as only the star of each point is stored. The details of the data structure are discussed and compared with other solutions currently available.

1. Introduction

New technologies such as airborne altimetric LiDAR (Light Detection and Ranging) or multi-beam echosounders permit us to collect millions – and even billions – of elevation points (samples) for a given area, and that very quickly and with great accuracy. One example of the use of that technology in the Netherlands is the efforts by Rijkswaterstaat to build an accurate digital terrain model (DTM) of the whole country containing as much as 10 samples per m^2 (the AHN2 dataset, www.ahn.nl). Such datasets have attracted a lot of attention because of all their possible applications, e.g. flood modelling, monitoring of dikes, forest mapping, generation of 3D city models, etc.

The main problems with massive point cloud datasets is that while they provide us with unprecedented precision, computers have great many problems dealing with very large datasets that exceed the capacity of their main memory. The result is that they cannot efficiently display all the information, let alone process them. Examples of point cloud processes useful for many applications are: derivation of slope/aspect, conversion to a grid format, control of double points, calculations of area/volumes, viewshed analysis, creation of simplified DTM, extraction of bassins, etc. Advances in technologies to collect data are basically superior to our ability to process data.

LiDAR or multi-beam datasets are formed by scattered points in 3D space, which are – in several cases – the samples of a surface that can be projected on the horizontal plane (a so called “2.5D surface”), see Figure 1a. This is of course not always the case but for several situations 2.5D is enough and it is therefore worth developing specific tools to handle this case since they can be highly optimised. While simply storing millions of unconnected points in a DBMS is no problem, the processing of LiDAR datasets needs more: we must be able to reconstruct the surface represented by the points, and we must also be able to access this surface to manipulate it (e.g. adding/removing new samples), and also to derive values from it. The surface gives us the spatial relationships between unconnected points in 3D space, which is required if processing on the surface is wanted. The best way to reconstruct such a surface is arguably with a triangulated irregular network (TIN) respecting the Delaunay criterion (Wang et al., 2001). As shown in Figure 1b, a TIN subdivides the space covered by the points into non-overlapping triangles, and these can be used to reconstruct the surface and answer neighbourhood queries (“which points are close to x and surround it?”). In Figure 1b, the vertex x has for instance 7 well-defined neighbours.



(a) (b)

Figure 1. (a) A 2D surface obtained from a set of points in the plane having a height as an attribute. (b) A triangulation of a set of points in the plane. The vertex v has 7 neighbours.

This paper investigates the use of a DBMS to store and manage not only the points coming from a very large LiDAR dataset, but also the TIN of these points. DBMSs are investigated since they are arguably the best tool to store and manage very large datasets (of any kind). The paper briefly discusses in Sections 2 and 3 current solutions for storing and processing point clouds, and then proposes in Section 4 an alternative which could have several benefits. This proposed solution goes beyond the usual "store points and edges and triangles" described in Section 2, and use recent advances in the compression of graphs.

2. Current solutions for storing and processing point clouds

Many companies offer solutions for the storage and processing of LiDAR data, and the problem is also being tackled in the academia. What follow is a short overview of the most interesting solutions, and does not claim to be a thorough review of all the possibilities. These were choosen because they are, as far as I know, the most used by practitioners and the tools that offer the most functionalities to handle massive point clouds.

2.1 Commercial software

The company **Terrasolid** offers several useful tools for the processing of LiDAR data. The main problem here is that, as is the case with many other tools, the size of your computer's main memory decides how many points can be processed. Basically, once the memory is full, transfer of data between the disk and the memory starts, and if this happens too much, the computations might simply stop (called trashing). The first solution that Terrasolid and others provide is tiling a big dataset into smaller parts that all fit into memory, and working on one given part at a time. While this solution is viable in some cases, for any processing that goes across a "seam" this can be problematic, e.g. calculation of areas/volumes of features, simplification often need a global view of the dataset, flow modelling, etc. The second solution is thinning, which means that when reading the input file, the algorithm will only read every 500th point for example. While it permits the user to visualise and process the data, it somehow destroys the idea of working with high-resolution altimetry data in the first place.

Since version 9.2, **ArcGIS** also provides a new type for the storage of very large TINs. In a nutshell, the Terrain type stores all the points in a DBMS, but the triangles are not explicitly stored. The user must select so-called "vertical indexes" so that a hierarchy of TINs is created. For each level in the hierarchy, ArcGIS selects representative points to form the TIN (the method used is clear, it appears that they are randomly selected). Because each TIN is small in size when compared to the original dataset, one can use the usual processing for TINs as available in ArcGIS (see Peng et al. (2004) for more details). However, this solution suffers from the same problem as Terrasolid, that is if the TIN created is larger than the memory there is no guarantee that the processing operations will terminate. For instance, if one wants to create a high-resolution grid from a TIN with all the input points, then if the TIN size (points + triangles + topological links) is too big then it will simply not work.

Oracle Spatial, since version 11g, also provides new types for the storage of LiDAR datasets: a Point Cloud type and a TIN type, which are similar. For building a Point Cloud, first the input points are "bucketed" into cells containing a given maximum of points (let us say 2500), then a spatial index (a R-tree) is constructed, and finally each point within a cell is also indexed (for TINs, the triangle are simply defined by a reference to 3 vertices, and are also indexed within one cell). Preliminary tests at GDMC/TUDelft have shown that many intermediate results are stored to build the spatial indexes, and these use a lot of memory (Tijssen, 2009). Also, the TIN type does not store the topological relationships between triangles (simply the 3 vertices), which means that these types are probably more targeted towards the storage of very large datasets, and that the manipulation of the surface is limited (to the best of my knowledge Oracle Spatial does not offer any functions to manipulate or analyse the surface at this moment).

2.2 Academia

To deal with massive datasets, one can also design external memory algorithms (Vitter, 2001). These basically use disks to store temporarily files that do not fit in memory, and instead of using the mechanism of the operating system, design explicit rules for the swapping between the disk and the memory. These are an alternative to using a DBMS, and have been used for a few LiDAR processing problems, see for instance Agarwal et al. (2006) and Arge et al. (2006). The drawback is that the design of such algorithms is rather complex, and for different problems different solutions have to be designed.

An alternative solution is the streaming approach of Isenburg et al. (2006a, 2006b), which mixes ideas from external memory algorithms with different ways to keep the memory footprint very low. The input points are processed in a certain order, and removed from memory when they are not needed anymore. The idea was applied to create a billion-triangle TIN, and they succeeded in under one hour with a conventional laptop (Isenburg et al., 2006b), which improved by a factor of 12 the fastest existing, that of Agarwal et al. (2005). The streaming ideas are very useful for certain local problems (e.g. interpolation), but unfortunately cannot be used (or it would be extremely challenging) for global processes such as simplification or flow modelling.

3. Storing TINs in a DBMS

The simplest way to store a TIN in a DBMS is to define a new datatype Triangle which contains either the coordinates or the IDs of three points in the point cloud (a polygon could also be used). As mentioned previously, that permits us to store the triangle, but topological relationships between the triangles have to be computed (an expensive operation) if analysis of the surface is wished, and it required spatially indexing massive amount of triangles.
An alternative is to implement in a DBMS the triangle-based data structure used by most triangulation libraries, e.g. that of CGAL (Boissonnat et al., 2002). As shown in Figure 2, it considers the triangle its atom and stores each triangle with three pointers to its vertices and three pointers to its adjacent triangles. The vertices and edges of each triangle must be oriented consistently (e.g. counterclockwise) if we want to manipulate or perform some operations on a triangulation. This structure stores topological relationships between triangles, but increases the storage space required and at least two tables (points + triangles) have to be maintained when the structure is updated.

More complex data structures for planar subdivisions could be also used (e.g. the DCEL (Muller and Preparata, 1978), the half-edge (Mäntylä, 1988)) or the Generalized Maps (G-Maps) (Bertrand et al., 1993, Lienhardt, 1994)), but they are very verbose (several primitives are stored and these are linked by several pointers) and would most likely not be optimal in a DBMS.

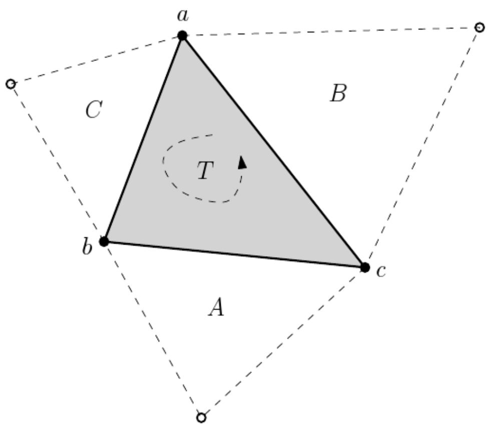


Figure 2. With the triangle-based data structure, the triangle T has 3 pointers to its 3 vertices a, b and c; and to its 3 neighbouring triangles A, B and C. The pair vertex-triangle is organised in such a way that for instance a is 'opposite' to A.

4. A star-based data structure

The approach I propose in this paper to store and process massive point clouds is storing the whole TIN in a DBMS (the points, the triangles and their topological relationships), and rely on the DBMS for memory management. Since the whole TIN is available (and does not need to be recomputed), one can readily query it and manipulate it. The processing operations would then be built over the database without designing special memory management methods. This solution will most likely be a bit slower than solutions working with main memory, but once the data structure is working and efficient, the benefits of a DBMS should compensate for the lack of speed.

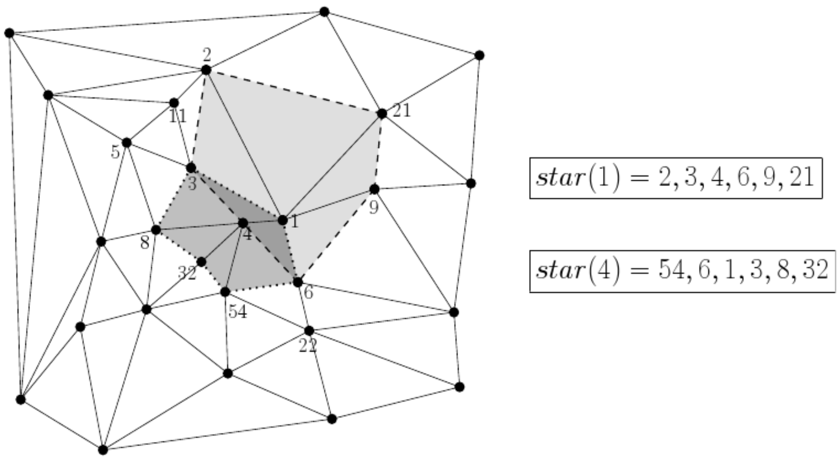


Figure 3. A triangulation with the stars of vertices 1 (light grey with dashed lines) and 4 (darker grey with dotted lines) shown.

The data structure proposed uses recent advances in the compression of graphs, particularly the "star-based" structure of Blandford et al. (2005), which was developed for main memory. It does not store triangles explicitly, instead the star of every vertex is stored. As shown in Figure 3, the star of a given vertex v, denoted star(v), consists of all the triangles that contain v; it forms a star-shaped polygon. The star of a vertex can be simply represented as an ordered list of the IDs of the vertices on the star; a triangle is formed by the centre of the star plus 2 consecutive vertices in the star list. If the star of every vertex in a dataset is stored (thus each star overlaps several other stars), then we can implicitly store not only the triangles of the TIN, but also all their topological relationships. Adjacency between triangles are stored, but also incidence relationships between vertices and edges and triangles (so one can for instance navigate counterclockwise through all the triangles incident to a given vertex). Observe that each triangle is present in exactly 3 stars and each edge in 2 stars (in Figure 3 the edge (1, 3) is represented in the stars of vertices 4 and 2, in the opposite direction). The data structure is therefore also akin to the the half-edge (Mäntylä, 1988) where each edge is stored twice, one for each direction.

ID	x	y	z	star
1	3.21	5.23	2.11	2–44–55–61–23
2	5.19	29.01	4.55	7–98–111–233–222
3	22.43	15.99	8.19	99–101–73–23
...	...	...	...	...
5674	221.19	15.23	37.81	309–802–793–1111

Figure 4. One example of a table to store the star of each point.

As shown in Figure 4, implementing this data structure in a DBMS is straightforward: only one table with 5 attributes (id, x, y, z, start) is needed; star is an array of IDs. Perhaps the main advantage of such a structure is that the use of a separate spatial index is not necessary. Indeed, we can rely on the topological relationships between the triangles to perform typical queries such as: (i) given a location (x, y), which triangle contains it? ; (ii) range queries: return all the triangles contained in a rectangle; (iii) obtain the triangles adjacent to a given one; etc. Functions have to be implemented inside the DBMS (server-side programming), and only a standard index on the ID column is needed because several SQL queries will be have to be made. An example of such a function is the walking in the triangulation (see Figure 5 for an example) where adjacency relationships are used to navigate inside a TIN (see Mücke et al. (1999) for a detailed explanation). Using SQL queries to perform such a walk is usually problematic, but with server-side functions this is possible.

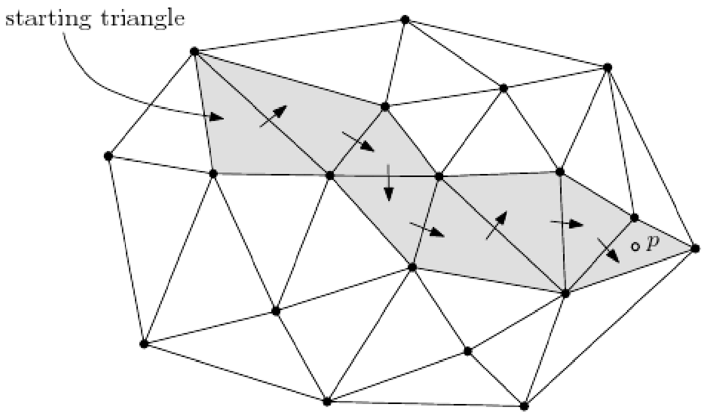


Figure 5. The WALK algorithm for a triangulation.

5. Discussion

The proposed star-based data structure is significantly different from what is usually available in a DBMS (i.e. each element (point and/or triangle) is stored independently and a spatial index is used for queries). It permits us to bypass the use of a spatial index at the cost of possible slower answers to queries. However, it is space efficient compared to other solutions described in this paper and the topological relationships between the triangles are explicitly stored, which permits us to process and manipulate the structure. With solutions where the triangles are stored independently, the reconstruction of these relationships is necessary each time an operation needs to be performed.

Furthermore, the structure permits dynamic updates (addition and removal of triangles and/or vertices are possible, with local updates (Blandford et al., 2005)), and these could be implemented directly in the DBMS. The GISt group at the Delft University of Technologies is currently working on implementing and optimising this structure, with basic functions such as interpolation in TINs, slope/aspect derivation and conversion to grids.

The solution proposed is valid not for 2.5D models but also for any boundary representations that can be triangulation (also for closed volumes), and the ideas are readily extensible to higher dimensions. As a consequence, the idea of storing stars of edges in 3D would permit us to efficiently store tetrahedra, and would offer an alternative to the structure of Penninga (2008).

References

P. K. Agarwal, L. Arge, and K. Yi. I/O-efficient construction of constrained Delaunay triangulation. In *Proceedings 13th European Symposium on Algorithms*, pages 355-366, 2005.

P. K. Agarwal, L. Arge, and A. Danner. From point cloud to grid DEM: A scalable approach. In A. Reidl, W. Kainz, and G. Elmes, editors, *Progress in Spatial Data Handling – 12th International Symposium on Spatial Data Handling*. Springer, 2006.

L. Arge, A. Danner, H. Haverkort, and N. Zeh. I/O-efficient hierarchical watershed decomposition of grid terrain models. In A. Reidl, W. Kainz, and G. Elmes, editors, *Progress in Spatial Data Handling – 12th International Symposium on Spatial Data Handling,* pages 825–844. Springer-Verlag, 2006.

Y. Bertrand, J. F. Dufourd, J. Françon, and Pascal Lienhardt. Algebraic specification and development in geometric modeling. In *Proceedings TAPSOFT'93*, volume 668 of *Lecture Notes in Computer Science*, pages 75–89, Orsay, France, 1993.

Daniel K. Blandford, Guy E. Blelloch, David E. Cardoze, and Clemens Kadow. Compact representations of simplicial meshes in two and three dimensions. *International Journal of Computational Geometry and Applications,* 15(1): 3-24, 2005.

Jean-Daniel Boissonnat, Olivier Devillers, Sylvain Pion, Monique Teillaud, and Mariette Yvinec. Triangulations in CGAL. *Computational Geometry – Theory and Applications*, 22: 5-19, 2002.

Martin Isenburg, Yuanxin Liu, Jonathan Richard Shewchuk, and Jack Snoeyink. Streaming computation of Delaunay triangulations. *ACM Transactions on Graphics*, 25(3): 1049-1056, 2006a.

Martin Isenburg, Yuanxin Liu, Jonathan Richard Shewchuk, Jack Snoeyink, and Tim Thirion. Generating raster DEM from mass points via TIN streaming. In *Geographic Information Science – GIScience 2006*, volume 4197 of *Lecture Notes in Computer Science,* pages 186-198, Münster, Germany, 2006b.

Pascal Lienhardt. N-dimensional generalized combinatorial maps and cellular quasi-manifolds. *International Journal of Computational Geometry and Applications*, 4(3): 275-324, 1994.

Martti Mäntylä. *An introduction to solid modeling*. Computer Science Press, New York, USA, 1988.

Ernst P. Mücke, Isaac Saias, and Binhai Zhu. Fast randomized point location without preprocessing in two- and three-dimensional Delaunay triangulations. *Computational Geometry – Theory and Applications,* 12: 63–83, 1999.

D. E. Muller and Franco P. Preparata. Finding the intersection of two convex polyhedra. *Theoretical Computer Science*, 7: 217–236, 1978.

Wanning Peng, Dragan Petrovic, and Clayton Crawford. Handling large terrain data in GIS. In *ISPRS 2004—XXth Congress*, volume IV, pages 281–286, Istanbul, Turkey, 2004.

Friso Penninga. *3D Topography: A Simplicial Complex-based Solution in a Spatial DBMS*. PhD thesis, Delft University of Technology, Delft, the Netherlands, 2008.

Theo Tijssen. Puntenwolken in Orable RDBMS (in dutch). Internal report, GDMC, Delft University of Technology, 2009.

Jeffrey Scott Vitter. External memory algorithms and data structures: dealing with massive data. *ACM Computing Surveys*, 33(2): 209–271, 2001.

Kai Wang, Chor-pang Lo, George A. Brook, and Hamid R. Arabnia. Comparison of existing triangulation methods for regularly and irregularly spaced height fields. *International Journal of Geographical Information Science*, 15(8): 743–762, 2001.

Marine high-density data management and visualization

M.A. Masry, P. Schwartzberg
CARIS Ltd., Canada TU

Abstract

Several different engineering disciplines make use of massive collections of point cloud data. Both LiDAR and multi-beam sonar systems, among others, generate these types of data. Because the points can be distributed randomly within a 3D volume, it can be difficult to spatially index, store and visualize this type of data. Furthermore, the relative novelty of point cloud data means that workflows for processing it have not yet been firmly established in the industry. CARIS has developed a robust and flexible system for managing massive point clouds, as well as a workflow for procesing them. The system can efficiently store and index well over 1 billion points in such a way that they can be queried and processed efficiently. The technology allows the stored points to be visualized interactively, even over a network, and is already integrated with our bathymetric data processing pipeline. This paper will present aspects of CARIS's point cloud storage system and explain its utility in a workflow for processing marine data.

1. Introduction

Modern sensor systems produce huge volumes of data. Sonar and LiDAR systems and other acquisition systems generate large quantities of unorganized 3D points, each of which can have multiple attributes and can be treated as an independent location in 3D space. This paper is intended to provide an overview of the way in which this data type is used in the bathymetric data processing workflow and, in particular, the new point cloud storage system used by CARIS software applications to store and process this data.

Bathymetry is typically measured using multibeam systems – large, powerful sonar systems that cover large areas (swaths), and are typically deployed using survey patterns that cause their measurements to overlap. Modern multibeam systems are capable of generating vast quantities of data: a Kongsberg Simrad EM3002 single head system, for example, can record 40 samples per second, with 254 beams per sample. The sonar signals are typically processed by an online bottom-detection algorithm that returns only those signals that represent the intersection between the sonar pulse and the seafloor.

Sonar returns are typically stored in terms of range and angle. These pairs are then processed using software such as CARIS HIPS that combines them with time-stamped position, heading and attitude data recorded from GPS and motion sensors. The output is a series of multi-attributed, georeferenced 3D points. The horizontal positions are typically recorded in terms of latitude and longitude, and the vertical position in terms of depth. The attributes stored with each point often include measurements of the uncertainty of the point's horizontal and vertical locations. A typical, week-long shallow water survey can record well over 256 million 3D points, and surveys that produce well over 1 billion points are not uncommon. Storing and visualizing this collection of data is challenging.

In this paper we describe CARIS's system for handling high-volume point clouds, provide some visualization examples, and discuss our future goals for this work.

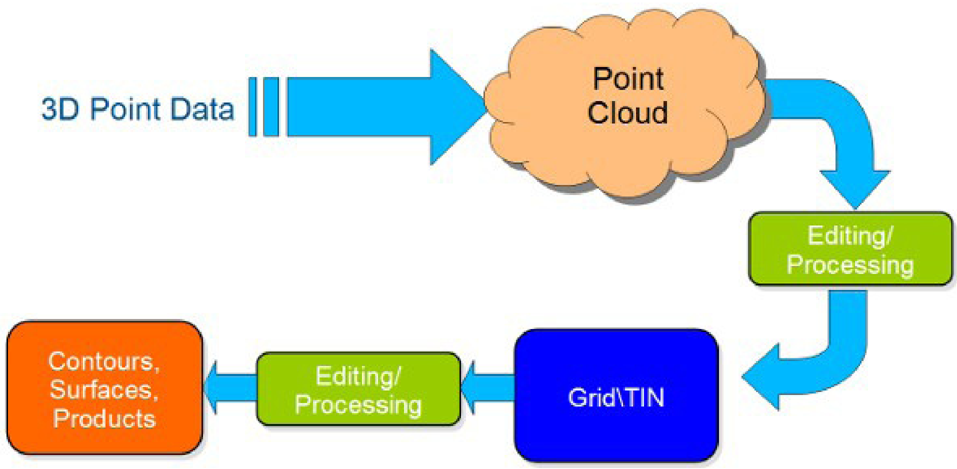


Figure 1. A typical bathymetric point post-processing workflow.

2. Downstream processing

The generation of georeferenced 3D point data from raw sonar data and time-stamped sensor information is largely automated, though user intervention is often required to correct poor input parametrizations or sensor noise. The next-generation 3D Point Cloud storage system described below, allows for sophisticated querying and visualization that can aid in this process. Three-dimensional points can be imported into either format directly from open file formats such as GSF, LAS, or any number of formatted text files, as well as through the above mentioned processing workflow for sonar data.

Bathymetric data processing typically focuses on the generation of 2.5 dimensional surfaces from processed 3D points. These surfaces can either be regular grids or Triangulated Irregular Networks (TINs). While they may be delivered as a final product, these often form the inputs for a range of subsequent processes. These processes include the contouring and sound selection processes that form an important part of the production process for nautical charts, as well as volume computations, interpolation tools and surface generalization algorithms. A typical post-processing workflow is shown in Figure 1.

3. The CARIS 3D point cloud

CARIS has implemented a sophisticated 3D point cloud data structure that can store billions of multi-attributed 3D points. The point cloud is the current data storage mechanism for 3D point data in CARIS Bathy DataBASE and will be incorporated into CARIS HIPS in a future release.

3.1 Data structure

Each point in the point cloud is stored using 64 bit double precision values, and can carry with it multiple attributes and flags; the attributes are grouped into bands and stored independently from the positional data to minimize I/O accesses, while the flags are stored along with the positional information. While the point cloud supports independent points, several points can also be grouped together into a single “multiple return” data point. The point cloud is bounded by a 3D spatial volume that is partitioned into multiple overlapping sub-volumes during the creation process. The points within each sub-volume are also ordered in a way that facilitates efficient visualization.

To create the point cloud, an insertion program analyzes an incoming stream of points to determine their aggregate spatial volume, sub-divides the volume appropriately, then determines which points lie within each sub-volume. The inserter can split each sub-volume recursively if the number of points in the sub-volume exceeds a given tolerance. The inserter also structures the points in such a way that new points can be added to the point cloud post-creation.

The points in each sub-volume are written to a storage device using the CARIS Spatial Archive (CSAR) framework. A diagram of a typical technology stack developed on top of the CSAR framework is shown in Figure 2. The framework provides developers with a set of tools that can be used to handle large volumes of multi-dimensional data by partitioning it into pieces called "chunks", each of which is given a unique key that can be used to retrieve it from storage devices such as standalone files or relational databases. The CSAR framework provides CARIS with a device-independent way to read, write and cache large data volumes in memory; it is used as part of the infrastructure of the CARIS Bathy DataBASE, allowing users to visualize and process data stored in the database directly, without having to first import the data.

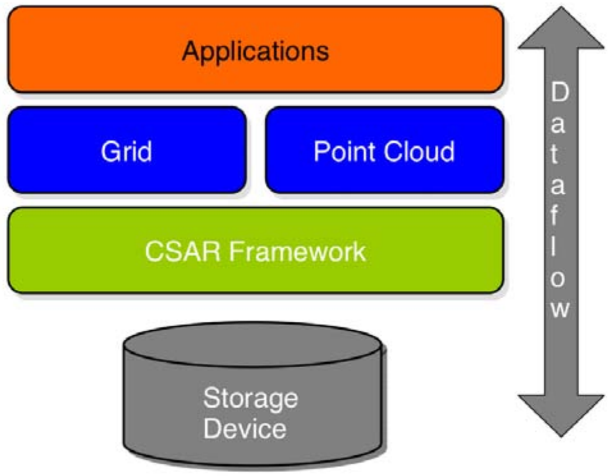


Figure 2. The application technology stack.

Chunks can be of different types, though in practice the chunks in the point cloud are homogeneous in type. Chunks are first created in a memory cache, then written to disk as the cache begins to fill. Each chunk passes through an I/O layer that serializes it into a contiguous memory block using a predefined serialization scheme, and optionally compresses it to minimize the size on disk. Because it is important that point locations be returned exactly, the system only implements lossless compression. When the chunk is read back from disk as part of a read call, its endian type is checked, then adjusted to the machine endian type if necessary. The chunk is then deserialized and loaded into the chunk cache, which may, depending on the amount of data currently loaded into the cache, cause another chunk to be written to disk to free up some available memory.

The point cloud creation process can specify the size of each chunk and the mapping of the points in each sub-volume into one or more chunks. Because I/O system performance generally varies with the number of I/O operations required to move data to and from storage, the number of points in each sub-volume generally has a significant effect on system performance, and must therefore be chosen carefully. The importer divides the stream spatially into sub-volumes using an octree-based volume subdivision scheme that begins with a tree of a single level. When the number of points in a sub-volume reaches a limit, it is split into 8 sub-volumes at the next lower level. This process continues until the number of points in a sub-volume is such that it can be stored in a single chunk.

The point cloud importer has been tested on point sets of larger than 1 billion multi-attributed 3D points, all of which were successfully imported into a single CARIS point cloud; preliminary results suggest that much higher numbers of points can be stored without further modifications to the structure.

3.2 Visualization and interaction

Structured indexing and spatial query mechanisms were implemented to allow subsets of the clouds (both points and attribution) to be retrieved from storage with a minimal number of I/O operations. The query mechanisms can return the subset of points within a constrained 3D volume, 2D projection, attribute range, flag setting, or resolution constraint. Once the 3D points have been imported into the point cloud structure, it is often necessary to select a subset for removal or editing. The CARIS Subset point editing utility in both CARIS HIPS and Bathy DataBASE uses these query mechanisms to select a subset of the point data from the point cloud storage formats and allow users to modify it interactively. The utility can be used to identify outliers, modify point attributes and flag points for exclusion from later downstream post-processing.

The sub-volumes in the point cloud are organized hierarchically to facilitate rapid 2D and 3D visualization and interaction with well over a billion points: overviews can be drawn in less than a second. Each point can be coloured using the value of any attribute, allowing for rapid and in-depth investigation of trends in the data. Figure 3 shows an example of use of different attributes to colour a dataset from the 2008 Shallow Survey sample set. Because only points within a specified visual tolerance are loaded onto screen, the visualization engine can sustain frame rates of 20-30 frames per second on a Pentium Core 2 Duo notebook computer with 256MB of discrete graphics memory.

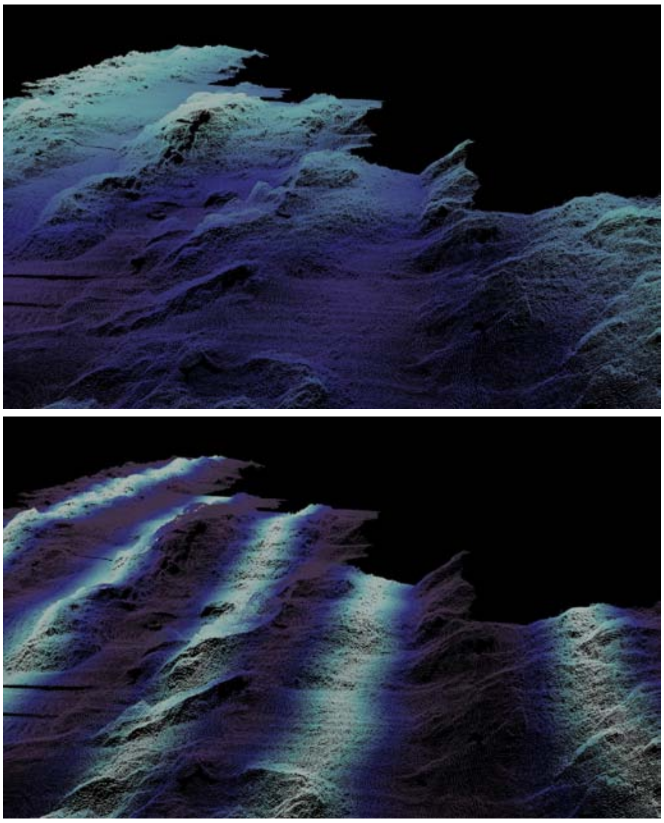


Figure 3. Two views of part of a 3D point cloud containing several million points. The top view shows the data colored by depth, while the bottom view shows the data colored by horizontal uncertainty. This data was taken from the Shallow Survey 2008 sample data set (http://shallow survey2008.org/).

Points from nearby sub-volumes are loaded into the 3D scene using a background thread, so the 3D visualization system maintains interactive frame rates while manoeuvring around the point volume, even when the data set size exceeds the size of main memory. Data can be loaded into the 3D scene at a rate of 5 – 10 sub-volumes per second, though this rate depends greatly on the performance of the underlying I/O subsystems; fewer sub-volumes per second can be loaded into memory over a network connection than from a local disk, for example.

4. Conclusions and future work

This paper presented a typical workflow for marine bathymetric point data acquired from multibeam sonar systems, and CARIS's 3D point cloud system, which is capable of storing very high volume sets of unorganized 3D point data. Several important issues remain to be addressed, however. The amount of computation required to create a point cloud remains high and, ideally, it would be possible to create and process a point cloud of several billion points in real-time while preserving those aspects the structure that facilitate rapid query and visualization. Furthermore, because of the sheer volume of data, editing the points in the cloud is a challenge, so the development of structures that allow for rapid modification of collections of points within the cloud, or the addition of new points to the cloud will be important. Finally, the increasing volume of points generated by today's sensor systems will pose scalability challenges: compressing the point data, and improving the I/O efficiency and cache performance of point cloud data structures will help address this issue.

References

[1] Wimmer, M and Scheiblauer, C, Instant Points: Fast Rendering of Unprocessed Point Clouds, Proceedings of Europgrahics Symposium on Point-Based Rendering (2006), pp 129-136.

[2] Wand, M., Berner, A., Bokeloh, M., Jenke, P., Fleck, A., Hoffmann, M., Maier, B., Staneker, D., Schilling, A., Seidel, H.P., Processing and Interactive Editing of huge point clouds from 3D Scanners, Computers and Graphics 32, 2 (2008), pp 204-220.

[3] Rusinkieweicz S, Levoy, M., Qsplat: A Multiresolution point rendering system for large meshes, Proceedings of ACM SIGGRAPH (2000), pp 343-352.

Scalable visualization of massive point clouds

G. de Haan
TU Delft, Computer Graphics & CAD/CAM group, the Netherlands

Abstract

Graphical renderings of "raw" point clouds can be visually attractive. When combined, the many individual points and their attributes convey spatial structures, especially when viewed in a fluent 3D fly-through and on a stereoscopic display. For such interactive visualization, a sufficiently responsive graphics update is essential. Where many software packages support simple rendering of smaller (subsets of) point clouds, the size of point clouds currently acquired easily surpasses their rendering capabilities, even on a modern graphics card. We addressed this issue in recent experiments on the visual exploration of LiDAR (Light Detection And Ranging) datasets in our Virtual Reality systems. Out-of-core terrain rendering techniques are applied, which originate from graphic rendering engines used in games and flight-simulators. In this document some of these techniques and results are highlighted, and challenges in balancing rendering performance versus visual quality are discussed.

1. Introduction

Acquisition of real-world scenes using LiDAR (Light Detection And Ranging) or other 3D scanning technologies produces enormous quantities of 3D data, typically in the form of large, unstructured point clouds. The 3D visualization of this raw data and its derivative features is often necessary to perform visual (quality-) inspection and to perform interactive selection and manipulation operations in 3D. But foremost, a 3D visualization empowers the human visual perception; it provides a recognizable and intuitive spatial context and is essential for the presentation of data to a non-expert audience.

Datasets of sizes up to some hundred megabytes (several millions of points) can be readily visualized using standard tools on commodity hardware. However, even scans of rather small environments can easily lead to large data files consisting of billions of individual points. An attempt to render these datasets with standard tools often result in low, non-interactive frame-rate or simply exceed the machine's limits in terms of main memory and graphics memory. In addition, it is in many cases necessary to have the data reside on a central server and only allow visualization from remote machines.

For these reasons, a *scalable* approach for visualization should be able to optimize the visual quality and interactivity given the limitations of the available rendering and network resources. To achieve this, powerful *out-of-core* data structures and rendering algorithms are needed. We consider the following three main scalability challenges of an interactive point cloud visualization system: *rendering, visualization and interaction*.

For the *rendering* part, a special data structure has to be designed that not only allows for high-quality real-time rendering but also supports editing operations. The rendering subsystem has to be scalable in terms of available processing power and main memory and the data structure should support methods for streaming the content to remote clients. For providing insight in all the properties and characteristics of the data, *visualization* techniques need to be developed that are able to visually represent the various features present in the data, e.g. comparison of two datasets in time. Finally, human-machine *interaction* issues need to be addressed. That is, the user has to be able to intuitively navigate, select, annotate and manipulate parts of the 3D data.

The current work in progress on these challenges is discussed in the remainder of this document. First, the architecture of our current prototype visualization system is described in Section 2 and 3. Here, some basic methods and data structures for scalability in terrain rendering are explained. Sections 4, 5 and 6, describe the integration of point clouds in our visualization system and the visual properties. We demonstrate some of the current results on visual quality and on performance in various point cloud datasets in Section 7. In Section 8, we highlight some related work and systems in the area of scalable rendering and point cloud visualization. Finally, in Section 9, we discuss current limitations and new challenges of current systems and indicate possible future directions of research and development.

2. Current visualization system

A schematic overview of the flow of data and its processes is shown in Figure 1.

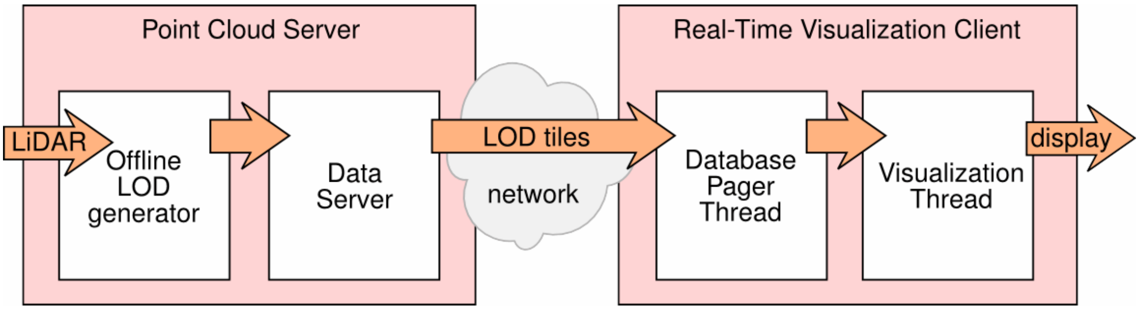


Figure 1: Point cloud visualization pipeline overview. The LiDAR files are first processed into LOD (Level-of-Detail) tiles before they can be used by the system.

For visualization we use “VRMeer”, an open source software toolkit developed at the TU Delft Computer Graphics Group. It is used to prototype and to develop interactive Virtual Reality applications. It builds on features of the OpenSceneGraph (OSG) 3D graphics rendering toolkit. This is a modern *scene graph* engine and accompanying 3D toolkit that is widely used, under active development and open-source. The VRMeer toolkit adds support for various 3D display systems and 3D input devices and basic 3D interaction techniques. Although the base programming language is C++ for both OSG and VRMeer, we constructed wrapping layers for the Python programming language for both. This allows for a more high-level prototyping of new functionality and applications. As many other software toolkits and libraries for GIS data, such as the Geospatial Data Abstraction Library (GDAL), come with Python interfaces, this facilitates fast integration and linking of general 3D visualization applications with GIS specific data formats.

3. From point clouds to terrain rendering

Techniques for point cloud visualization have many parallels to terrain rendering techniques which are used in modern games and high-performance flight-simulators (e.g. SGI OpenGL Performer). OpenSceneGraph also provides the functionality for rendering large terrain models. These real-time, out-of-core terrain rendering techniques are created by combining level-of-detail (LOD) scene graph nodes and the paging of data in- and out of memory from disk or network. In this section a brief outline is given of the functionality of this mechanism.

In an OSG-based visualization client, the *visualization thread* determines which parts of the scene graph is visible, and enables for each node the suitable LOD level for this viewing distance. If this LOD level is not yet available in memory (e.g. it resides on disk), the *Database pager thread* is requested to load this data. Without visibly pausing the application, it will start loading the data from directly disk or from network in the background and insert the graphical

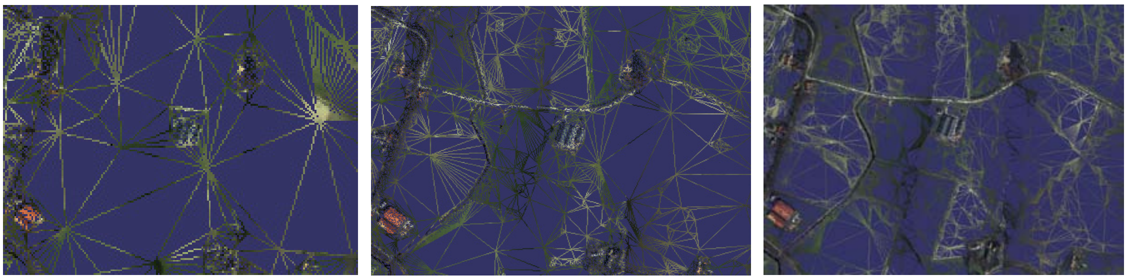


Figure 2: Screenshots in wireframe mode of the same area in the AHN2 terrain as seen from a large viewing distance (left, scaled) through to a small distance (right). As the viewing distance decreases, a different level-of-detail is visible with more accurate triangulation and higher resolution texture information.

node in the scene graph when it becomes available. At the same time, those nodes that are not visible in the scene are scheduled for removal from system memory.

OSG is accompanied by VirtualPlanetBuilder (VPB), a software tool to generate these terrain models. This program intelligently combines (several tiles of) Digital Elevation Maps (DEMs) and textures into a Triangulated Irregular Network (TIN) model at several levels of detail in a pre-processing phase. The final model contains scene graph nodes, each of which in turn contains terrain nodes with discrete level-of-detail tiles as its child-nodes. We used VPB to process AHN grid data (generated from filtered point clouds) which is available from the AHN test set. This set consists of several tiles in the ESRI ASC format (e.g. gridded DEM with 0.5m sized grid cells, height value in meters for each cell) and raw RGB data acquired by the LiDAR scanner in a geo-referenced GeoTIFF format. This data can be fed directly to VPB, whereas the necessary format conversion and coordinate system transitions are performed by the GDAL toolkit. As a result, VPB generates binary OSG files (IVE format). Several quality and terrain settings can be selected. Depending on the amount of tiles and required quality of the results, generation may take between several minutes to several days.

The resulting terrain model can be loaded directly in any OSG model viewer or our specialized, VRMeer based terrain viewer, see Figure 2. The visual quality differs with the distance the viewer takes to the landscape, and the spatial frequency of landscape features. A consistent frame rate of 60 frames per second is achievable during 3D exploration, without disturbing artefacts when loading in different levels of detail.

4. From terrains to point cloud rendering

When using a DEM without filtering (c.f. AHN2 unfiltered), buildings appear to be “draped” with a sheet, see Figure 3 (left). This is caused by the interpolation between neighbouring grid cells. The results are most visible in areas where the continuity of the terrain is disturbed by structures on top of the terrain, such as buildings and trees. However, raw LiDAR measurements contain a wealth of 3D information, of which much is lost in the post-processing steps that are used to generate DEMs. In the AHN case, a point cloud consists of approximately 20 million points with XYZ coordinates (floating point number, RD coordinate system, cm notation, ~3 cm accuracy). In our work we aimed to directly visualize these point clouds without removing information or introducing new sampling artefacts. When this point cloud is directly converted to a scene graph model by using a graphical point primitive for each point, it provides a rich visual scene with different visual characteristics, see Figure 4 (right). However, frame rates will drop far below interactive rates, e.g. to one frame per second. To still maintain sufficient interactive performance, the point clouds need to be pre-processed in special scalable data structures, in a similar fashion to the TIN models as described above.

Figure 3: Visual comparison of two visualization techniques of the same area of the AHN2 dataset. The triangulated mesh (left) is generated from the unfiltered, gridded DEM and suffers from "draping effect", most visible in trees and buildings. The point cloud (right) is generated from the unfiltered point cloud and more directly depicts LiDAR measurements. It does suffer from see-through effects and "holes", most visible around and underneath buildings.

5. Data structures

We have developed a point cloud pre-processing mechanism which generates OpenSceneGraph data structures for point clouds. These data structures can be paged into memory and allow control of the levels-of-detail. In this section, some of the characteristics of this processing and data structure are discussed.

In a similar fashion to what the VirtualPlanetBuilder does for TINs, the pre-processing program constructs *a hierarchical tree structure* in an OSG scene graph, and *distributes* the available points of the point cloud over the nodes at different levels in this tree. For each node in the tree containing a subset of the points, it's LOD distance is set. When viewed from a distance, the level-of-detail algorithm displays only the higher level leaves, so only a subset of points in the tree are visible. A careful *balancing* of number of points and the number of levels of detail is necessary to avoid loading and rendering too much data.

First, we select a hierarchical data structure for spatial subdivision depending on the characteristics of the point cloud. We currently support a regular binary tree, a quad-tree or an octree. For large, terrain-like datasets a quad-tree is the most efficient of these options, whereas for a detailed scan of a single building an octree would be more suitable. The depth of the tree is currently estimated based on the density of points in the source LiDAR file, and the expected detailed viewing distance. We would like to balance the number of points in each node to minimize the overhead of the hierarchical structure, and not have too many points in a single node. Based on the domain size and tree depth, the hierarchical data structure is created in an OSG scene graph and the spatial extents of the individual bounding boxes are generated.

Second, for each point in a point cloud, we determine in which spatial tile it should be placed in. Here, in contrast to many other LOD approaches, *points are not duplicated* in order to save memory and bandwidth capacity. This approach was selected because we can render complementary point clouds that completely overlay each other, we consider this a special feature of point clouds. As a result, we not only place points in the leaf nodes of the LOD tree but also in all of the branch nodes.

To determine how high in the LOD hierarchy a point should reside, a *sampling strategy* is used. The goal of a sampling algorithm is to have a representative subset of the points in each of the LOD levels. As a result, if only the lower LOD level (largest distance) is rendered, a "reasonable" impression of the overall measured structure should be obtained. Once the LOD level of the point concerned is determined, it is placed in a point set in the corresponding LOD level for that spatial subsection of the domain.

We currently use one very basic sampling mechanism that uses a bitwise interleaving scheme on the explicit ordering of the points in the point cloud. For example, every 16th point is placed on the lowest LOD level, every other 8th point is placed on the second lowest level, until the remaining points are placed on the correct, higher and more detailed LOD levels. If the measurement process is a complete random sampling of the structure and unordered, this approaches a random sub selection. It can be calculated what the expected number of points and density will be in the resulting LOD tiles. In the discussion section other suitable alternatives to this approach are discussed.

Per individual point we can also store available attributes. Examples of these are measured point color or intensity. We can also use ortho-photos to color individual points by performing a color look-up for each point in the corresponding photograph at its x and y location. A large downside to per-point attributes is the increase in data-size, e.g. a 3-float RGB component adds 12 bits per point, doubling the data size with respect to only position encoding.

6. Visualization

As soon as the point cloud tiles are loaded in the scene graph, they are visualized on-screen. It can be rendered concurrently with other (geo-) data and interface elements.

The visual representation of the point samples is done through rendering OpenGL point primitives. Only a limited perception of depth can be obtained from a single 3D point. However, when many points are grouped, visual depth cues of perspective, visual compression, motion parallax and occlusion give a compelling spatial impression.

The drawback of point clouds is that they do not necessarily give a continuous or filled visual representation, e.g. no sampling gaps are filled. This is not a problem when viewing from a distance, but can be if observed from close by. This impression may be improved when a point is rendered larger than single pixel. For this, a graphics *shader program* can determine the virtual size of a point to generate a larger (or smaller) *splat* on-screen, based on the distance to the viewer. This improves the depth perception of individual points, especially when they are closer to the viewer as they scale with motion and occlude other points and may fill holes. Also a *fogging* component is included to simulate *haze*, which emphasizes distance by decreasing the saturation of colors when points are more in the distance. If no color is present in the data, pseudo coloring can be used, e.g. to map a point's height to a color value, e.g. see Figure 4.

A combined representation of both raw point clouds and a terrain model generated from filtered data may provide a reasonable impression of a landscape with sharp features such as buildings. However, we did observe disturbing effects which appears to be a visual shadow or ghost. It suffers from see-through effects and “holes” in the point cloud, through which the same color is also seen on the ground surface. This is mostly visible around and underneath buildings, an could be solved by removing colors on the ground surface when buildings are present. We can also integrate 3D surface models of buildings that were modeled separately or that were reconstructed from the point cloud.

7. Resulting interactive visualizations

The techniques described in the previous section were applied for the interactive visual exploration of several LiDAR datasets in our Virtual Reality systems and demonstrated on numerous occasions. In addition to regular desktop visualizations, we can use *stereoscopic* rendering and 3D interaction in our Virtual Reality systems. The stereoscopic rendering can dramatically emphasize depth perception for complex point clouds which are not dense or lack structural or color information features.

General 3D viewers often do not include good interaction techniques to study 3D topography data. Google Earth is an example of a specialized viewer for topography content. However, this application has a main drawback that it cannot simply load custom terrain models and/or point clouds. The main aspect of this viewer is the 3D navigation. We created simplified mouse based view navigation, as well as navigation with specialized devices such as the *Space Mouse* and the *Wii Balance Board*. On more immersive VR systems, such as a Workbench or CAVE system, one can also use head tracking and 3D pointing devices. By using our Python-based prototyping system, the application can be quickly reprogrammed to include case-specific viewing techniques, specialized interfaces, or loading of models. As an example, we included mouse-based sketching to quickly draw sketches on a 3D terrain for communicating details.

Figure 4 depicts the use of a point cloud based visualization integrated with a flooding simulation. Although the simulation used here is non-realistic and simplistic, the visual effect and the 3D context of the information within the point cloud is found compelling. We are currently planning to integrate realistic simulation results from running the SOBEK Flooding Simulation from Deltares, and in the future to allow interactive steering of these simulation runs.

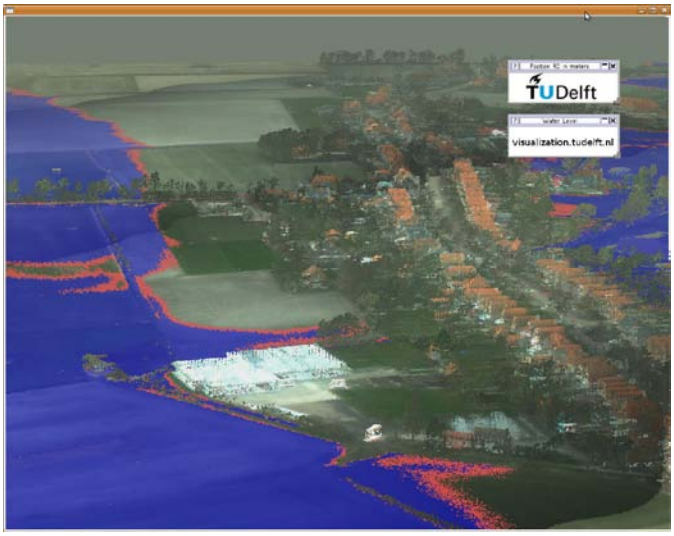


Figure 4: Screenshot of a point cloud -based visualization integrated with a simplistic flooding simulation. The use of per-point color, fogging, and larger point-sizes increases the perception of depth and context of simulation results.

8. State-of-the art: beyond scalable rendering to scalable visualization

Various techniques have been presented for out-of-core point cloud visualization, e.g. Kreylos et al. (2008a) and Wand et al. (2008) make use of a multi-resolution spatial data structure. In a pre-processing step the raw data is analyzed and written to disk in a special format that allows for efficient reloading of higher resolution parts of the scene. Both methods achieve interactive visualization of raw data and were demonstrated with sizes of up to 63 GB. The system presented by Kreylos et al. (2008a) allows for simple editing operations whereas the XGRT system presented by Wand et al. (2008) is a proof-of-concept of a complete tool chain for 3D data processing. In both systems editing of the scene that leads to reasonably small changes in the overall scene structure can be performed in real-time. A different method for extremely fast and efficient visualization of terrain data based on a regular grid is presented by Bhattacharjee et al.

Figure 5: Screenshot of the AHN2 dataset (filtered buildings, no color) surrounding TU Delft campus, ~6 tiles in view at 20 fps @ 1680x1050 (mono), nVidia Geforce 280 GTX. Performance when viewing 100 tiles currently drops to 5 fps.

(2008). In this work, most parts of the rendering pipeline are outsourced to the graphics card (GPU) and high frame rates are achieved on small-to-medium sized data sets. The method is only suited for structured point clouds (height maps) and does not support out-of-core data.

In the field of computer graphics and visualization, much work addresses issues of performance of data handling, data processing and rendering. However, equally important are visualization techniques; these are responsible for adding abstract or derived features to the rendering. Especially with Point Cloud datasets, elements of uncertainty, variability and the detection of changes are of features interest that require visual representation, see also the work by Pauly et al. (2004) and Girardeau et al. (2005). In their work, Butkiewicz et al. (2007) demonstrate the possibilities of visual analysis for exploring change detection based on multiple LiDAR scans. Recent developments demonstrate various new ways of displaying feature and abstract information cross-linked with geo- information (such as point clouds), see Butkiewicz et al. (2008). In addition, the possibilities to interactively act on the data enable users to navigate, select, annotate and manipulate parts of the data. For example, Kreylos et al. (2008b) demonstrate the use of 3D interactive visualization for use in a geoscience scientific workflow. Here, large datasets are visually explored and potential hypotheses are generated, where the 3D perception and intuitive interaction with the data plays an important role. Fuchs et al. (2009) propose an interactive cycle of knowledge-based analysis and automatic hypothesis generation. In this work, starting from initial hypotheses, created with linking and brushing, the user steers a heuristic search algorithm to look for alternative or related hypotheses. The results are analyzed in information visualization views that are linked to the volume rendering. Individual properties as well as global aggregates are visually presented to provide insight into the most relevant aspects of the generated hypotheses.

9. Discussion and future directions

Our current approach as of now can be made scalable enough to visualize the full AHN2 set, under special circumstances that is. However, one currently needs to manually select several parameters for LOD tree depth and sampling interleaving. This first results in much manual tweaking and unnecessary processing of failed runs. In addition, it will not provide a scalable solution in terms of lower bandwidth and slow clients or the updating or editing of new information in tiles. It is under current investigation how the individual parameters influence performance and visual quality. We plan to provide more flexible approaches for sampling the data and develop metrics. An interesting approach would be to allow continuous Level-Of-Details and streaming of data, based on the performance of the network and the client. This should entail a better sampling strategy or defining an *ordering* in the importance of specific points.

Another topic of interest is the adaptive sampling strategy for the various levels of details. The current basic sampling strategy does not produce convincing results on some datasets. For example see Figure 6, where a dataset has regularity in point ordering, causing aliasing, or is not heterogeneous in terms of sampling density, causing holes or too dense areas. To the latter effect, for example, local point size could be increased when sampling density is low.

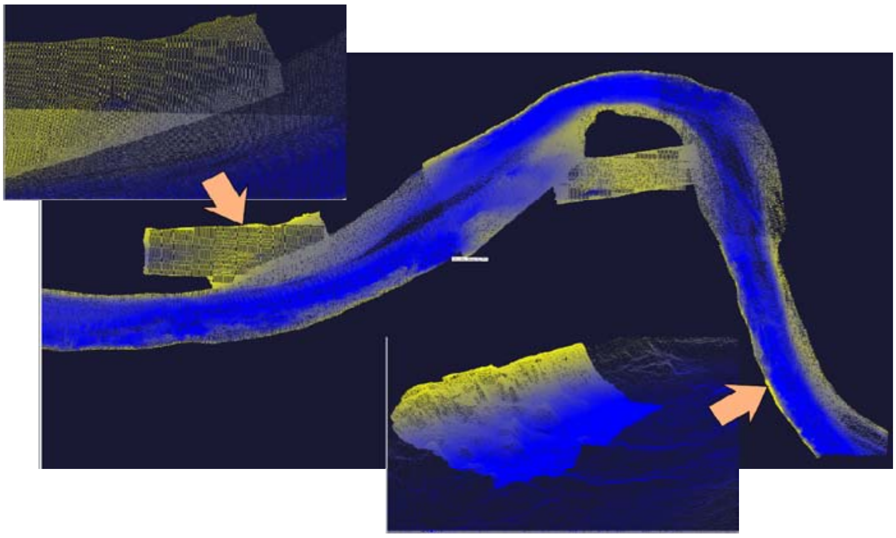


Figure 6: Point Cloud of a Bathymetry dataset. The inset top-left indicates aliasing errors caused by the regularity in data and the sampling algorithm, the other inset shows a region where point sampling density is much larger than its surrounding area.

An advantage of our OpenSceneGraph-based system over custom high-performance rendering approaches is that it benefits from compatibility, maturity and new developments in and surrounding the toolkit. For instance, currently new work is being done to integrate OpenGL ES support, which will enable small or embedded rendering clients such as mobile devices. Another interesting development is the integration of more mature GIS functionality within this framework, such as is being done through projects like OSGgis, OSGearth. An alternative development would be that current popular GIS applications adopt similar 3D viewing components in their existing applications.

In conclusion, the main objective of future work in this area is to both design fundamental algorithms and develop software assets which allow visualization and interaction on all the 3D data in a point cloud geo-database. More specifically, the objectives concern rendering (software methods and framework for scalable pre-processing and rendering), visualization (definition and evaluation of novel (human-supervised) feature extraction and feature visualization methods) and interaction (definition and evaluation of interaction techniques for 3D navigation, selection, annotation and manipulation).

Acknowledgements

The AHN2 dataset is copyright of RWS. The bathymetry dataset is copyright of RWS Zeeland.

References

Bhattacharjee, S., Patidar, S., Narayanan, P. J.: *"Real-Time Rendering and Manipulation of Large Terrains."* , ICVGIP, 2008.

Butkiewicz, T., Chang, R., Wartell, Z., Ribarsky, W., *"Visual Analysis and Semantic Exploration of Urban LIDAR Change Detection."*, Computer Graphics Forum, 2007.

Butkiewicz, T.; Dou, W.; Wartell, Z.; Ribarsky, W.; Chang, R.; *"Multi-Focused Geospatial Analysis Using Probes"*, IEEE Transactions on Visualization and Computer Graphics, Volume 14, Issue 6, Page(s): 1165 – 1172, Nov.-Dec., 2008.

Fuchs, R., Waser, J., Gröller, E., "*Visual Human+Machine Learning*", IEEE Visualisation, 2009.

Girardeau-Montaut, D., Roux, M., Marc, R., Thibault, G., *"Change detection on point cloud data acquired with a ground laser scanner"*, 2005.

de Haan, G., Griffith, E.J., Post, F.H., *“Using the Wii Balance Board(TM) as a Low-Cost VR Interaction Device”*, ACM Virtual Reality Software and Technology (VRST), 2008.

de Haan, G. and Post, F. H., *"StateStream: a developer-centric approach towards unifying interaction models and architecture"* In Proceedings of the 1st ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS), 2009.

Kreylos, O., Bawden, G. W., and Kellogg, L. H., *"Immersive Visualization and Analysis of LiDAR Data"*, In Proceedings of the 4th international Symposium on Advances, 2008.

Kreylos, O., Bawden, G., Bernardin, T., Billen, M. I., Cowgill, E. S., Gold, R. D., Hamann, B., Jadamec, M., Kellogg, L. H., Staadt, O. G., and Sumner, D. Y., "*Enabling Scientific Workflows in Virtual Reality"*, Proceedings of the 2006 ACM international Conference on Virtual Reality Continuum and Its Applications (VRCIA), 2006.

Pauly, M., Mitra, N.J. , Guibas, L.J. , *"Uncertainty and Variability in Point Cloud Surface Data"*, Eurographics Symposium on Point-Based Rendering, 2004.

Wand M., Berner A, Bokeloh, M. , P. Jenke, A. Fleck, M. Hoffmann, B. Maier, D. Staneker, A.Schilling, H.-P. Seidel: *"Processing and Interactive Editing of Huge Point Clouds from 3D Scanners"*, Computers and Graphics, 32(2), 204-220, 2008.

Web references

[OSG]	OpenSceneGraph, http://openscenegraph.org
[OSGEarth]	http://wush.net/trac/osgearth
[OSGGis]	http://wush.net/trac/osgearth

Mapping the world with LiDAR

S. Coppens
TomTom – Content Production Unit (Tele Atlas)/Engineering, the Netherlands

Abstract

To map the world TomTom/Tele Atlas is continuously looking for new source material to improve the quality of the geographical database, to extend the coverage more quickly, to keep it up to date and all this with an acceptable cost. Hereby point cloud data is becoming more and more a source of interest to reach those goals. Mapping reality means modelling hundreds and thousands types of features/attributes and in this paper an overview will be given where point clouds can become an added value and another part of the paper will highlight that cloud data is not a silver bullet and brings along some challenges if one want to industrialize the process and make it more cost efficient, which of course introduces new business opportunities and challenges academically.

1. Opportunities

Point cloud data can be used to create and maintain several features or attributes in a geographical database. The following subsections will give an overview of several feature/attribute groups where this data could and is being used.

1.1 Road geometry

An obvious domain where point clouds can be an added value is related to road geometry. It could help us in modelling new road geometry or help us in improving the absolute and relative accuracy. Although it must be said that so far this domain has not been studied in a large extend. Certainly the relationships between the point cloud accuracy, relative geometry accuracy and absolute geometry accuracy contains a set of interesting unexploded academic research topics.
Further more the point cloud data can also be used to capture and maintain new geometry related information which is crucial for Advanced Driving Assistance Systems (ADAS) such as road curvature, road slope and road width. Curvature and slope challenge our academics to come up with new algorithms and arithmetic paradigms to for example derive splines from the point clouds. Splines are functions defined piecewise by polynomials and are used frequently by road construction engineers to model road curvatures.

1.2 Road attribution

Once the road geometry has been captured one can start thinking about which attributes that could be captured or maintained with point cloud data. And without too much brainpower one can come up with examples such as pavement, lane information, tunnels and bridges.
The pavement includes the detection of different pavement structure or material such as asphalt, concrete, cobblestone, etc. But it can also look at the pavement conditions, so basically how good the road is maintained.

The lane information is again an information set that could be derived from the point cloud data. One could extract the number of lanes, the lane divider type (e.g. solid line, interrupted line, single, double, etc), the lane type (e.g. exit lane, parking lane, shoulder lane, etc) and even manoeuvre information based on the arrows on the road surface.
Another interesting and obvious feature one could extract from the point cloud data are the bridges and tunnels. And one can not only determine where they are but could also extract some attributes from it such as the length or the height of a tunnel. This information can of course not be derived from laser-alt but from terrestrial lasers, e.g. mounted on mobile mapping vans.

1.3 Road furniture

The next level we can use the LiDar for is any features or attributes along the road geometry. There is still a huge amount of interesting data along our roads that are very important to be captured within a geographical database. The list below gives an idea of features that could be captured via this data; some are more exotic, read challenging, then others. For all of them one can again detect the presence of the features but for some of them one can also extract attributes.

- *Side walks*. Determine the presence of a side walk but also attributes such as the width, pavement type, etc.
- *Curbs*. Determine the height of a curb.
- *Signs*. Determine the position of a sign but also the dimensions. And perhaps it is even feasible to categorize signs (e.g. sign post, speed restriction, stop sign, advertisement panel, etc.).
- *Trees*. Determine the position of a tree and also here perhaps the dimension.
- *Road lighting*. Determine the position and dimension information.

1.4 3D information

In the literature one can find a tremendous amount of material on how LiDar can be used to model buildings, to determine roof structures and of course also the height of a building. Point clouds can assist to build real 3D volume-tric objects and not only 2.5D (= building footprint with height attribute).
The buildings need to be also integrated within the environment and hereby is the alignment with for example road geometry instrumental, of course point clouds can assist here.
Finally the landscape should also be modelled in 3D and also here point clouds can help in creating a Digital Terrain Model (DTM) and assist in integrating the 3D buildings and road geometry within the DTM.

2. Challenges

When one has a geographical database which needs to fulfil different market segments one immediately gets in competing and conflicting requirements. Because each market segment has different priorities and focus, personal navigation is perhaps today very interested in road attribution, where in-car navigation can be more interested in ADAS related features and finally the online community is more into 3D City Models. Nevertheless they have one thing in common they all want to model the reality with a certain level of detail and each of them have there own challenges, personal navigation devices have a small screen, in-car navigation systems are not yet online and finally the online customers have still bandwidth restrictions.
Figure 1. Market - Focus and Importance shows some examples of the continuous clash between different markets and what those markets determine to be important and what they are focusing on.

All those requirements result in the geographical database getting following requirements:
- Should be at all time in line with reality.
  Model the reality correctly related to content integrity (= capture the right objects with the right attributes) and format integrity (= files/databases are technically correct).
- Should have different representations of reality.
  The same object can be modelled several times with a different level of detail. E.g. the number of polygons used to model a 3D building can have a big impact on the final representation of the building.
- The positional correctness should be very high.
  For a geographical database x, y, z coordinates are crucial. Every object should be georeferenced with high accuracy, can go from 5 m to 1m to even 10 cm absolute accuracy.
- The data size should be adaptable based on the requesting market.

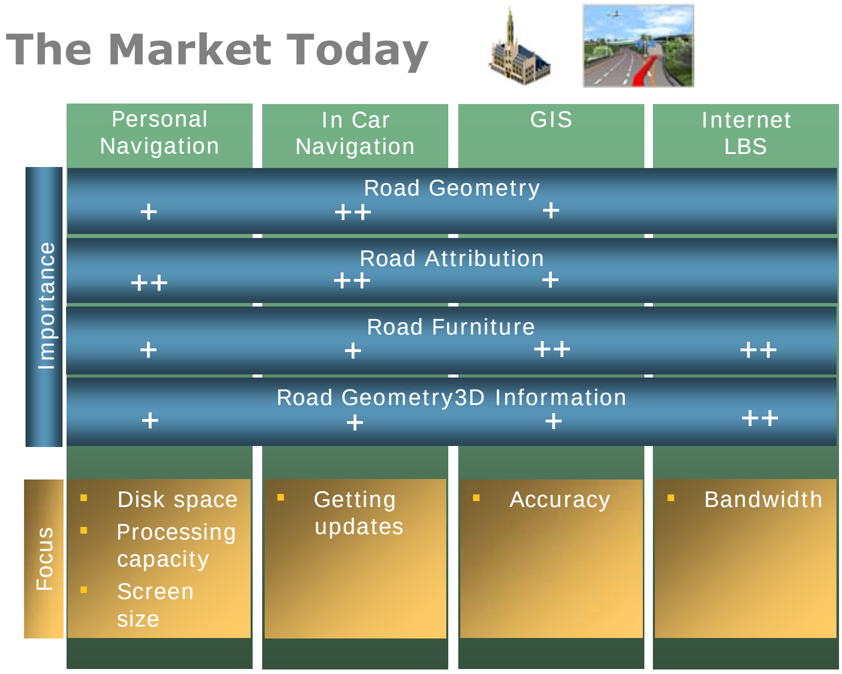


Figure 1. Market - Focus and Importance.

The opportunities section made clear that point clouds can help us in modelling the reality but if we now also take the four "simple" requirements, just listed above, into account then can we still say that point clouds can be a useful source material to create and maintain geographical databases?

Probably the last question posted just above can be answered positively but due to the fact the source material type is quite recent the production processes are quite immature and are a potential trigger of research projects.

2.1 Challenges to keep in line with reality

The main challenge with point clouds is the huge amount of data, which not only gives storage challenges but also processing/throughput challenges. Today it has been proven that one can use LiDar data to model 3D buildings but what if one has LiDar data for 40 million cities in the world and several capturing sessions of the same city? This increases of course complexity and mainly processing time tremendously. So how can we speed up processes?

- Can we extract in a more automated manner features from point clouds?
- Can we find differences between two deliveries covering the same object/area?
  Is it possible to perform change detection between two point clouds covering for example the same city? Can one have algorithms that detect the changes within a city based on those two point cloud deliveries?
- Can we combine different source materials to improve the efficiency?
- How to come to a smoother integration between DTM (Digital Terrain Model), buildings and road geometry?
- How to integrate buildings on a flat plane in the most efficient way, read as less as possible artefacts?
  Buildings on a hill will due to the slope not have a rectangle façade. However not all products have the DTM included and as a consequence one has a gap between the model and flat plane, Figure 2 illustrates this.

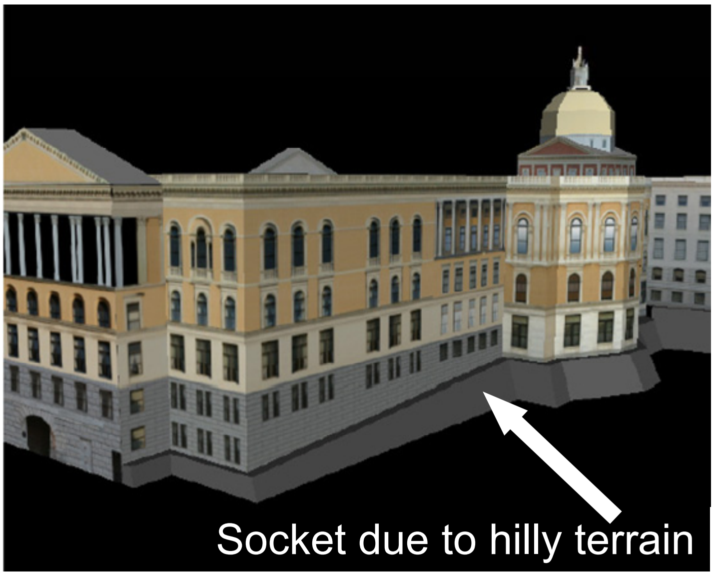


Figure 2. Socket example.

2.2 Challenges to have different representations of reality

In a lot of cases it is not a problem to have different representations of the same object. Although with the current technology it can't always be done in an automated manner.

- How can we derive less detailed representation of building from higher level of detail representation of the same building?
- How can the different models kept consistent with each other if adaptations are required?
- How can we reduce the data size for the same level of detail? Or what is the impact on data size between the different levels of detail? Can we apply any new lossless or lossy compression algorithms on LiDar data or derived features?

These challenges are mainly related to the 3D information.

2.3 Challenges to have high positional accuracy

Positional accuracy is becoming more and more important in applications, definitely for ADAS applications. Applications are using more and more the coordinates to adapt the application behaviour. For example orientation of the headlights of a car will change based on the road curvature data or the truck gear box is steered based on slope data. So we should have a clear view on how the accuracy of point clouds can have an impact on the positional accuracy of features in a geographical database.

- Are there processes to determine the positional accuracy of a point in the cloud?
  - What's the impact of reflection?
  - What's the impact of multi-path?
- How can the accuracy be improved of a point cloud?
  - Can the combination of different source material improve the quality?
  - What's the impact of combining several deliveries of the same object/area on the quality?

3. Conclusion

Point clouds are a quite new source material type which has a lot of potential to:

- Keep some features and attributes in a geographical database in line with reality. Hereby features/attributes come to mind related to road geometry, road attribution, road furniture and 3D information.
- Improve the positional accuracy of the geographical database.

Due to the fact the source material is new it brings along a number of challenges.

- Handling huge amount of cloud points.
- Extracting features/attributes from cloud points in an automated manner.
- Determining the positional accuracy of extracted feature/attributes from cloud point data.

As one can see point clouds can be helping us in modelling reality but brings along some new challenges for which we have to find answers. Nevertheless point clouds will help us in reaching our goals.

References

Bekiaris, E., Amditis, A., Panou, M.: DRIVABILITY: a new concept for modelling driving performance, Cognition Technology and Work, June 2003 (2003).

Berghout, L., Versteegt, E., van Arem, B.: Advanced driver assistance systems; Results of the State of the Art of ADASE-II (2003) available on http://www.adase2.net.

Carrea, P., Deregibus, E., Montanari, R.: The Invisible Hand: How to design the best interaction between Advanced Driver Assistance Systems and Car Users, SAE 2000-01- CO34 (2000).

IMechE (Institution of Mechanical Engineers): Advanced Driver Assistance Systems (ADAS): Vehicle Control for the Future (IMechE Seminar Publications) (1999), ISBN10 1860582435.

http://www.safespot-eu.org.

http://www.aide-eu.org.

Reconstructing 3d Buildings From LiDar Data: Ahmed F. Elaksher and James S. Bethel: School of Civil Engineering, Purdue University, 1284 Civil Engineering Building, West Lafayette, IN 47906, USA.

Topographic Laser Ranging and Scanning: Principles and Processing: Ji Shan and Charles K. Toth (Eds.): CRC Press: Boca Raton, FL. (2009). xvii and 590 pp., diagrams, maps, photos, images, index: ISBN 978-1-4200-5142-1.

Handling large amounts of multibeam data

S. Broen
Kongsberg Maritime, Hydrographic Department, Subsea, Norway

Abstract

Today's multibeam echo sounders are capable of outputting enormous amounts of data, at an exceedingly higher rate. Following the evolution of this technology, a need has arisen to be able to process and view this data before it is used to create an end-product. The user wants to be able to confirm that the echo sounder is performing as it should and that the data collected is making sense, before it is stored. He / she does not want to wait until the end of the day after the survey has completed to do this, or that day would have been wasted if the data collected does not meet expectations. It must happen in real time so the user can make necessary calibration and / or adjust the echo sounder settings before further data collection, if the user suspects that something is not right. Also, when such large datasets are being collected there is always a certain probability that individual data units are erroneous. Correcting such fluctuations manually can be very time-consuming so there is also a need to perform such corrections automatically.

1. Introduction

Kongsberg Maritime (KM) has developed technology to meet the multibeam echo sounder (MBE) users expectations of real time processing and display of depth points. It is being utilized through KM's software application for MBE control and data logging, named Seafloor Information System (SIS). A core technology in this application is named the Grid-Engine (GE). The GE will accept all depths in real time, in addition to seabed image data which is 3-5 times more, and make a bin model out of them. This bin model is then processed in real time, and a Display Model is created. This Display Model will create and maintain several grids in different resolution and detail. At the highest resolution, this Display Model presents the terrain model at full resolution. The operator can quickly change the map scale and get another view of the whole area in a lower resolution. The GE is primarily used to create terrain models from Multibeam data, but there is no restrictions so it has also been used to create terrain models from laser data. In this project the height of trees was displayed using the GE and the Display Model.

This text will initially give a rough introduction to how we collect data. Next it will explain the path that the data takes through the system, followed by an explanation of how data is processed and displayed. Finally there is a summary and some thoughts of what the future might bring.

2. Depth points

KM MBE systems consist of several core parts. Perhaps the most obvious parts are the transducers. There is a transmitter transducer and a receiver transducer, or an array of such installed underneath a boat vessel. The transmitting transducers convert electrical voltage to sound waves, and the receiver does the opposite. The received sound signal is basically sound echoes from the seafloor. Data are created based on the sound-travel time and received strength. Due to the sound-absorbing properties of the seafloor sediment, one is able to determine the sediment type based on the received signal strength. This is what constitutes the seabed image. The travel time of the sound pulse, combined with the known beam-emitting angle from the transducer, constitutes a basis for the depth points. This basis has to be combined with sensor data that monitor the vessel orientation. Finally an estimate of the water columns sound travel times has to adjust the points. After these steps of processing, we have achieved a best-possible (realtime) estimate of an actual point at the seafloor. Point clouds consist of consecutive swaths of depth points, as illustrated by Figure 1.

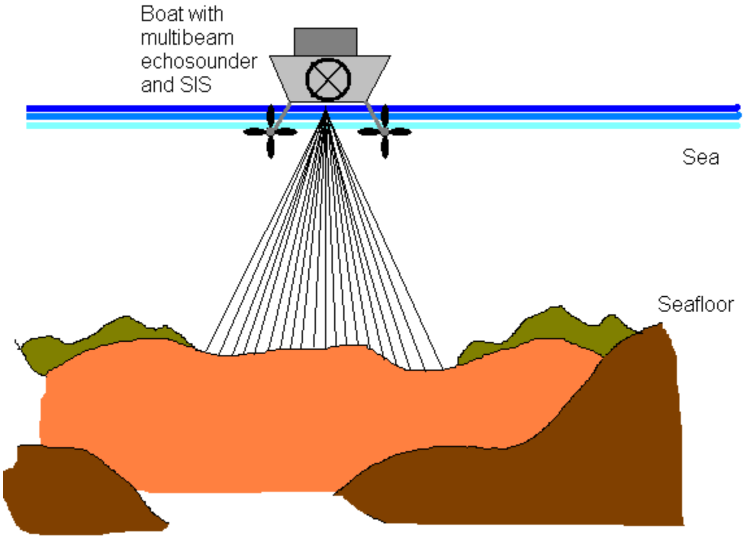


Figure 1. Collection of depth points.

2.1 Sensors

Necessary directional properties are being measured by a number of sensors. Examples of such properties are roll, pitch, heading and position of the boat, as illustrated by Figure 2.

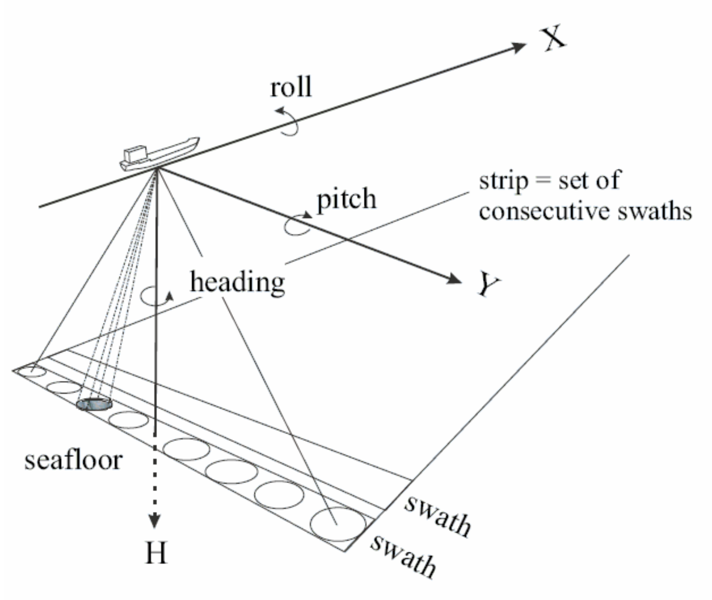


Figure 2. Vessel directional parameters.

Another important parameter is heave, which is also monitored by sensors. The data coming from these sensors are being combined with the depth data from the transducer in a Processing Unit (PU). The PU uses the sensor data to adjust the depth points, so that we can tell where the soundbeam actually hit the seafloor. Because of the often very long travel path of the sound waves, even the smallest change in vessel orientation will result in a corresponding change in depth points at the seafloor that are several multitudes larger. Because of this, it is critical that the sensors are being as accurate as possible. Also, it is necessary with a buffer to temporarily store raw depths because of the travel time of the beams. The received echo has to combine with the corresponding sensor reading at the time of transmittance. The result will be a seafloor point estimate, where we have assumed a consistent soundspeed through out the sea. Of course this is not the case, as we will see in section 2.2.

2.2 Water Sound Speed Profile

Snells law (*see Hovem, 2000; chapter 3*) states that a ray of sound traveling in liquid through the border between two layers of water with different densities and sound speeds, will result in a reflected and a refracted ray. The reflected ray will have an angle between its direction and border, that is the same as the angle between the incoming ray and the border. The refracted ray will have an angle based on the angle between the incoming ray and the border, and the difference in soundspeed and density between the layers. Because of this, a ray of sound moving from the echosounder to the seafloor will not be a straight line, but rather a line that will constantly have its direction altered at the transition between water layers. This can be seen in Figure 3, which is taken from SIS. The rightmost figure shows the zoomed-in area defined in the mid-figure.

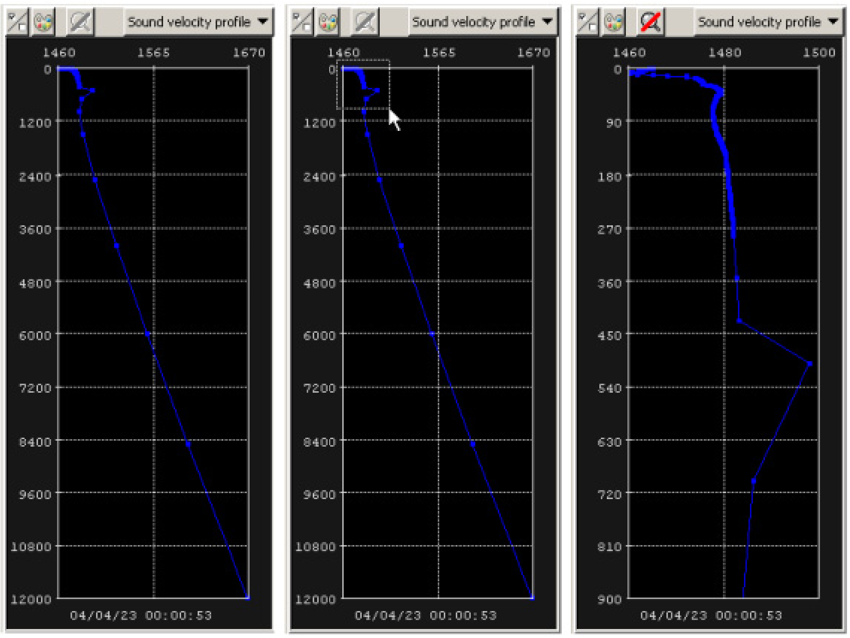


Figure 3. A ray of sound being altered by differences in water layers.

The smaller the angle between the border and the incoming ray, the larger the distance between measured consecutive points at the seafloor. In some cases, the outlying rays might never even reach the seafloor. Because of this property, we will need to measure the sound speeds in the ocean prior to surveying. To measure the sound speeds, it's common to use a soundspeed sensor probe that is sunk down the ocean, and then pulled up again. The result of this measurement will be a sound speed profile (SSP), and this profile will adjust the measured depth points.

2.3 Projection

We need to be able to reference our depths to the real world. To achieve this we need to **project** our depths onto a world reference system. There are various such systems, and they can be set be the SIS-user. When the necessary **parameters** in SIS are set correctly and some necessary pre-processing is done, the depth records may now be displayed and logged. The SIS geographical view may now look like Figure 4.

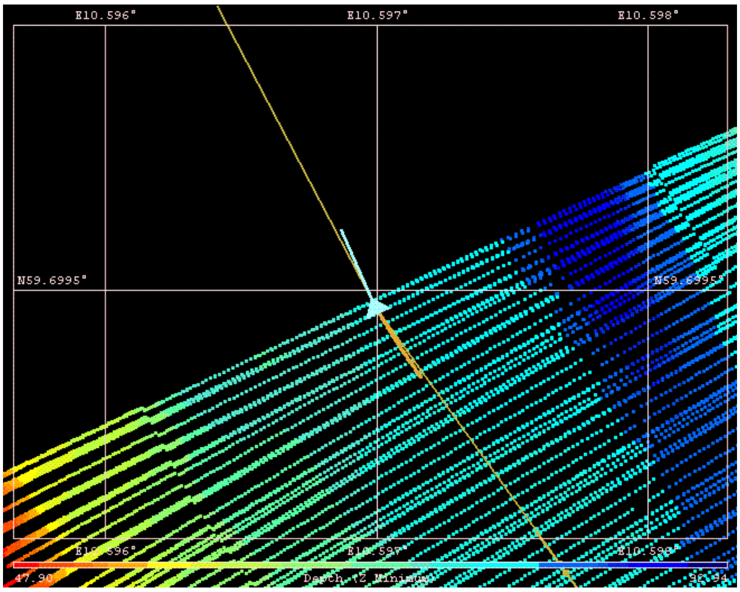


Figure 4. SIS Geographical View.

Note that every logged depth isn't necessarily displayed in the geographical view at once. The world coordinates relative to our selected projection can be seen on the left, right and upper part of the figure.

2.4 Datapath

The resulting data is then broadcasted on an Ethernet at which an operator PC (HWS) is also connected to. This PC has SIS installed on it. It is used to monitor data, log data and to control the echosounders. This datapath is illustrated by Figure 5.

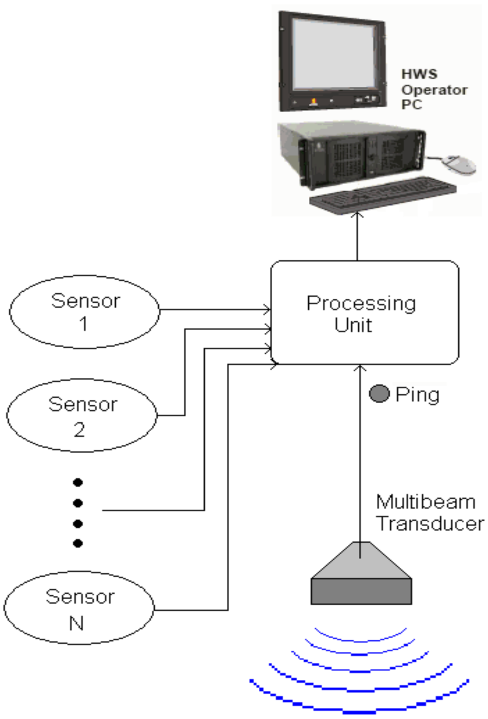


Figure 5. Depth point data path.

After just a few seconds with surveying the depth points start to seriously **stack up**. It becomes evident that there should be some mechanism to arrange all the points into something more eye-pleasing, like a terrain model. This is where the GE comes in.

3. The Grid Engine

The GE is a standalone server-based application. It is written in Java which ensures platform independence and automatic memory allocation/deallocation by the Java Runtime Environment. The bottleneck memory-wise then becomes the actual amount of installed RAM on the HWS. The server approach makes the GE see connected applications (SIS) as client ”websites”, and thus there are in principle no restrictions as to how many websites that can connect to this server. Storage of data also becomes more convenient this way.

3.1 GE datapath

The GE is split into several ***servlets*** which each handle different kinds of requests. SIS has a GE communication interface that responds to user input and transforms them to URL commands. For example, the user might want to zoom into a specific area, specified by a bounding area. Such an area can easily be drawn in the geographical view in SIS by left-clicking and move the mouse the draw a rectangle. An URL command is the sent from the GE-interface in SIS requesting the desired operation along with parameters, and taken care of by the *Survey servlet* in GE. This servlet may in turn take other servlets in use to perform its task. I will not go into each servlet in this talk, as it will be too time consuming. Figure 6 explains the basic flow of data inside a HWS. Figure 7 shows the basic flow of data of the GE.

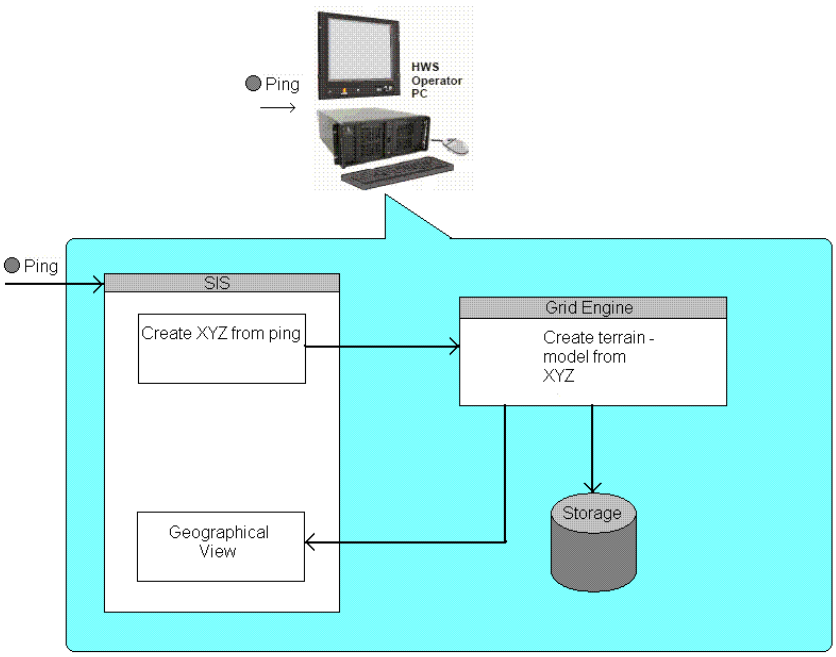


Figure 6. HWS datapath.

As can be seen in 7, data is being processed when it enters the system. A process grid is then created and stored to disk, then display grids are created and stored to disk. Connected clients may request to view the stored files. The following sections will explain each individual step of Figure 7.

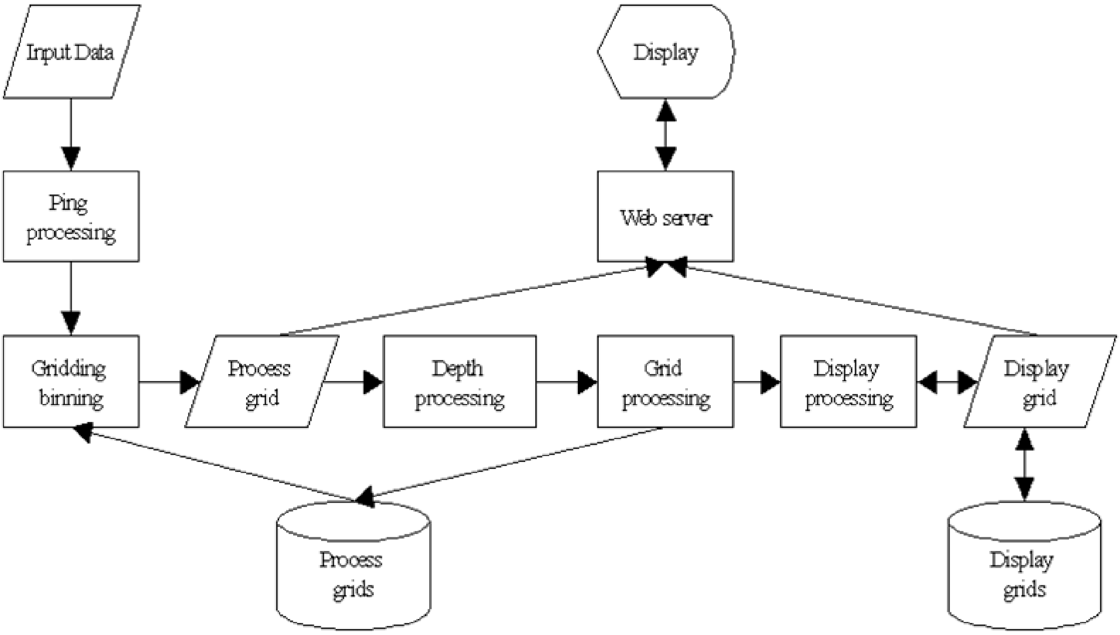


Figure 7. GE datapath.

3.2 Realtime ping processing

Realtime ping processing is performed right after the data enters the system (Røang et al., 2002), and before gridding and binning of data. The purpose of the ping processing is to flag points that probably are erroneous according to some relatively simple rules.

The rules can be divided into two main types; the first kind only looks at a single ping at a time, while the second kind needs a neighborhood around each beam in every ping. The rules are applied in sequence, and the single ping rules are always applied before the neighborhood rules.

The extraction of neighborhoods requires a buffer of pings, which results in a small delay. The size of the delay depends on how many pings needs to be stored in the buffer, which in turn is depth dependent.

There are two single-ping rules. First, the time rule, which flags beams whose two-way time range is longer than the time difference between two consecutive pings. This rule is rarely used. Second, the overhang rule, which flags beams that create an overhang situation where the loca-

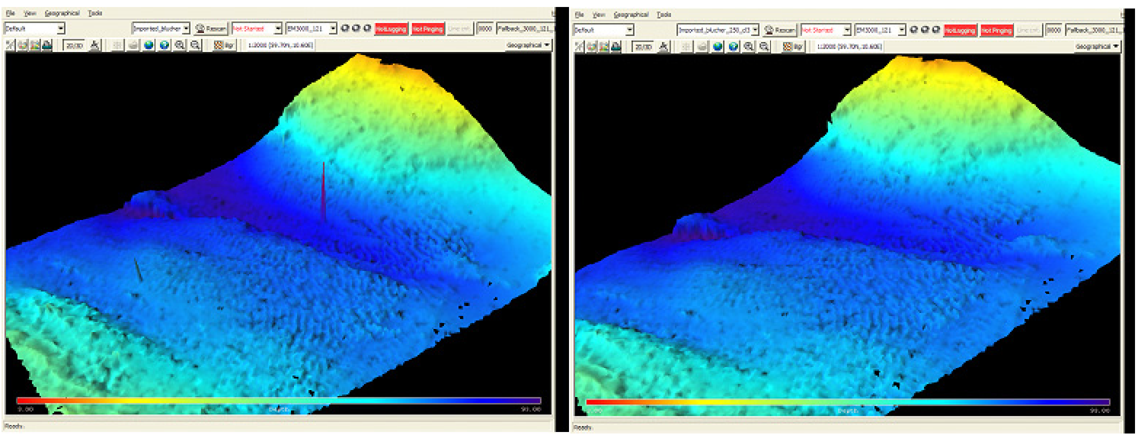


Figure 8. Removing spikes with ping processing.

tion of the depth of that beam is closer to the vertical projected depth from the sounder to the seafloor, than the depth of the neighboring beam that should have been closer.

There are two neighborhood rules. First, the median rule, which compares the absolute value of the depth difference between the center point and each point in the neighborhood. The center point is flagged if its depth differs too much from the neighbors. Second, the quantile rule, which is more adaptive than the median rule, in that it takes into account the local variability of the topography and the noise. Figure 8 demonstrates the result of a ping processing.

3.3 Realtime depth processing

Data are now being put into a Process Grid, where they can be processed realtime (Debese, 1998). A Process Grid contains a square number of *cells*. The number- and size of the cells (in m^2) may be set by the user. In deeper waters, larger cells are more appropriate in order to get required amount of points inside each cell. Points are sorted into the cells by their xy-value. When we later come back to the same area, data from the previous lines will be read back and reprocessed together with data from the next line. This is called Area Based Data Cleaning (ABDC). But before we can do this we have to split and merge all cells in the grid to ***processing units.*** Based on the number of points and the depth variation, we try to merge or split with neighboring cell(s). Like this, steeper seafloor will require smaller units in the xy plane. Units can also be empty. After an iteration of this sorting, the grid may look for instance like Figure 9.

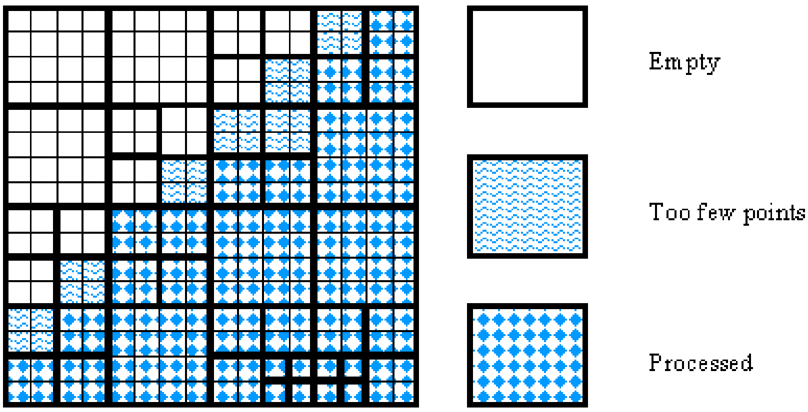


Figure 9. Arrangement of process units.

After the processing grid has been arranged into ***processing units***, the list of units is traversed and a surface estimate is calculated for each of them. The traditional approach is to simply use the mean depth value, meaning the surface approach will always be flat. Instead of doing this, the GE uses the iterated reweighted least squares (***IRLS) algorithm*** to calculate a ***first, second or third degree polynome surface*** that model the terrain much better (Debese and Bisquay, 1999). See the Figure 10.

The IRLS will take the set of points, and start off by calculating a ”least squares” surface. Now each point is given a ”weight” based on their distance to this surface. For each iteration, the surface is recalculated based on the point weights. Point weights will pull the surface closer, ”heavier” points have more influence than ”lighter” points. Then each point is given a new weight based on their new distance to the surface. This goes on until either the surface converges, or the number of iterations has passed a threshold. Note that outlying points (4 red dots) in Figure 10 have been flagged out based on their weight.

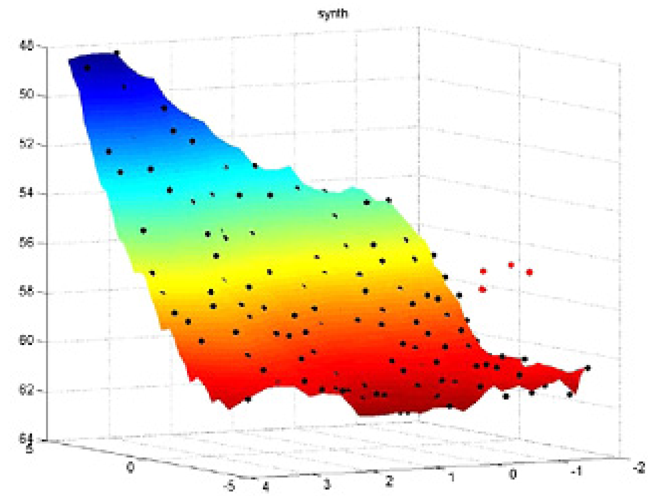


Figure 10. Terrain model estimated for one unit.

For each iterations, points are given new weights based on their residue according to a Tukey function, see Figure 11. This function has been subject to a lot of research, and the approach Kongsberg Maritime has chosen is based on Debese (1998).

Points that are (most likely) erroneous will be given little or no weight, and after the iterations they can be flagged out based on , for example, a threshold weight value. The valid points are sorted and some of the values from the minimum, maximum and median points are transferred to the ***display grid***. All "point set variables" (i.e. all the "z" values) are sorted and transferred separately to the display grid.

The display grid contains "point set" data and cell data. Point set data are minimum, maximum and median point data for each point set variable ("z "variable). Cell data are scalar data valid for the entire cell such as average backscatter and maximum depth difference. Data is stored in different "layers" (files on disk) where each layer represents a level of detail (LOD). SIS chooses which LOD to display based on the current span in latitude / longitude in the geographical view (zooming). The cell data for the lowest level are calculated in the processing grid using all valid points in a cell. When data are processed from one LOD to a higher LOD, data from one to four Display Cells are used to compute the values for the Display Cell above, see

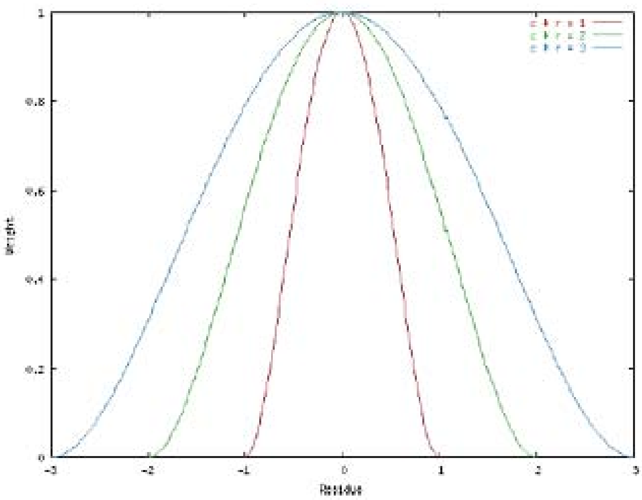


Figure 11. Tukey weight functions of residue and weight.

Figure 12. The minimum and maximum point set data is computed by taking the point set with the minimum and maximum value from the level below. The new median depth is computed by taking the median point set from the level below which is closest to the center.

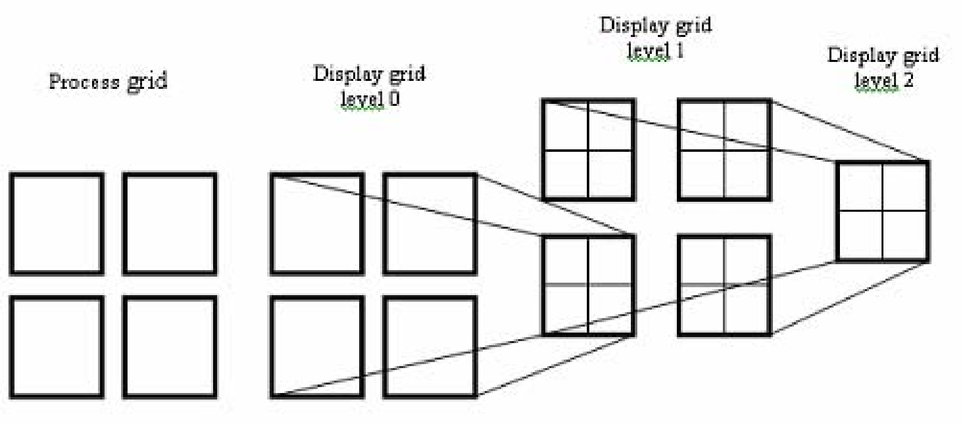


Figure 12. Process grids and display grids.

For the different point set data, separate min, max and median values are calculated for cells at each LOD. Cell data is either calculated as average, area average, minimum or maximum values of the values from the level below. (When area average is computed, the value is weighted with the number of process cells for that cell.).

The number of level of detail is dependant of the number of cells in the processing grid. The more number of cells in each direction the more level of detail. A processing grid with 64x64 cells will have 6 levels of detail, i.e. log2(64).

All the grids at all level of details are stored to disk and updated realtime. By request from SIS, the GE will read the files, and return them to SIS. This means that the user can access the grids in real time, and SIS may switch between the levels of detail seamlessly, based on how far in or out the user is zooming into the seafloor. Figure 13 shows an example of how the geographical view in SIS can display data at a given LOD.

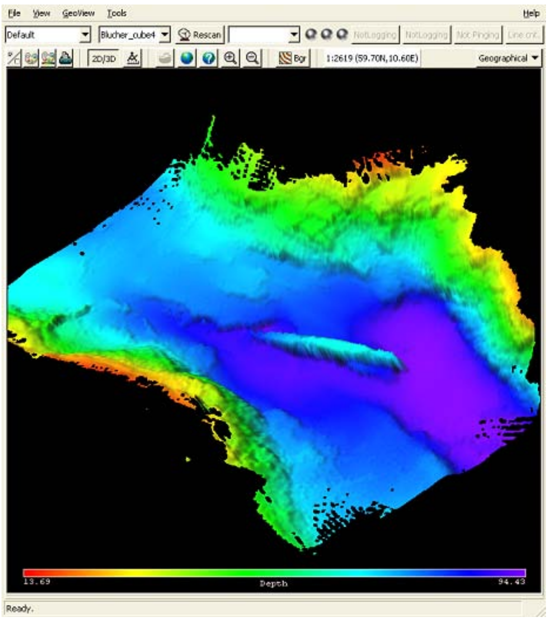


Figure 13. A terrain model made from depth points.

4. Summary and future developments

The KM MBEs produce massive collections of depths. These collections constitute our Point Clouds. The Point Clouds are sent to the GE, which process all data, and creates a terrain model at multiple resolution levels. The GE is under constant development, and has seen numerous upgrades since its release. The next version will contain functionality that removes the requirement that the SIS user manually has to set GE parameters (number of cells, cell size). The *dynamic* grid engine will figure this out on its own, and it will be featured in SIS in the near future. There are plans to further improve it in the future. For example, there are plans to make more effective memory allocations based on predicted vessel heading. So the GE will try to ”guess” which grids it must load into memory.

References

Debese N. (1998). Application d'un estimateur robuste à la détection des erreurs ponctuelles dans les données multifaisceaux: l'estimateur de Tukey. Rapport d'étude du SHOM, 002/98. Brest: SHOM.

Debese, N. and Bisquay, H. (1999). Automatic Detection of Erroneous Soundings in Multibeam Data through a Robust Estimator. Proceedings of the HYDRO ’99 conference, Plymouth, 8 pp.

Hovem, J.M. (2000). Marin Akustikk. Lecture syllabus, Institutt for teleteknikk/akustikk, NTNU, Trondheim, Norway.

Røang K., Eliassen I.K. and Heggelund Y. (2002). Real Time Data Cleaning of Multibeam Echosounder Data. Christian Michelsen Research AS, Bergen, Norway.

Virtualising large digital terrain models

G.G. Spoelstra
ATLIS, the Netherlands

Abstract

Chart producing agencies like Hydrographic Offices face major challenges today in keeping up with the ever-growing amount of data that is produced by modern echo sounders. At the same time users are demanding products faster and yet more reliable to keep pace with the more powerful and accurate systems they use. To support Hydrographic Offices and other chart producers, new innovative technologies need to be developed that will bring a great deal of efficiency into the current survey to chart work processes ensuring better products and a much faster time to market. Above challenges can be met by the introduction of V*irtual Continuous Model (VCM)* technology. The ATLIS SENS (Spatial Enterprise Nautical Solutions) Bathymetry product is utilising this technology in such a way that data only has to be stored once; all users use the same data and can define their own terrain models without the need of copying the often large volumes of data to an individual workspace. The number of models (both up-to-date and historical) is virtually unlimited thus providing a maximum of flexibility to expert organisations. Using SENS Bathymetry, large digital terrain models are virtualised by just storing model definitions based on available meta data instead of redundant storage of the data itself. The underlying Oracle Spatial technology that holds the archive of survey data ensures fast retrieval of seamless models that can be used for a wide variety of applications. Using proven database technology is a great advantage over technologies that are based on traditional file systems which are reaching their limits on today's operating systems.

1. Introduction

Hydrographic offices face increasing pressure to deliver products faster, yet more reliable than ever. With source data coming in at growing speed and volume, bathymetric data management is changing into a major bottleneck between data acquisition and chart production. At the same time users are demanding products faster and yet more reliable. Organisations that manage bathymetric data often use the data for various reasons: safe navigation, morphology, off-shore planning to name a few. The challenge with managing large volumes of bathymetric data is to keep everybody happy without the need to (re)build and manage models for each of these communities. A nautical cartographer requires a navigational safe model that is as up-to-date as possible whilst a marine morphologist might be interested in a series of historical models to analyse sediment transport. The problem gets even more complex as most organisations are under the obligation to archive their bathymetric models for liability reasons. In addition they also have to make data available as part of national and international data infrastructures.

Mandatory ECDIS (Electronic Chart Display and Information System) carriage requirement, announced by the International Maritime Organisation (IMO) will force ship owners into investments that are easier to accept when the Electronic Navigational Charts fuelling ECDIS provide a much richer and more up-to-date content than today's paper charts. Content that is derived from seamless databases avoiding the traditional paper chart patchwork and offering high density bathymetric information to the end user in a smooth and unambiguous way. The size and draft of modern vessels make it almost impossible to bring these ships into ports without the use electronic chart systems showing go/no go areas with sub-meter detail. Demand of contour intervals of less than a meter in electronic charts is therefore no longer an exception but require advanced algorithms and high performing technology.

In section 2, this paper will look into the changing demands for high density bathymetric data and explains why today's production processes of hydrographic organisation form major bottlenecks when it comes to building large digital terrain models compiled from many different sources. The associated challenge is illustrated by describing work of marine morphologists and the liability issued faced by National Hydrographic Offices. The concept of the Virtual Continues Model as a solution is discussed in section 3. Section 4 highlights some future requirements for further improvement of handling massive point clouds using modern data base technology.

2. Changing demands

2.1 Different user needs

Bathymetric information is collected for various reasons. Safe navigation is probably the most prominent one but dredging, environmental protection, morphology, engineering, off-shore explorations, archaeology, fishery and coastal zone management are other examples of applications that require a detailed and up-to-date picture of the seabed. Satisfying all these users requires careful planning of the bathymetric workflow. Laying pipe lines using bathymetric data prepared for safe navigation is not a wise thing to do, investigating environmental impact at a world scale using high density multi beam surveys a very costly and time consuming exercise.

2.2 New demands for bathymetric data

Most organisations responsible for collecting bathymetric data usually performed their task with a relatively narrow scope. Hydrographic offices are responsible for navigational charts, survey companies performing route surveys only collect their data to check the status of the pipeline or cable. Introduction and growing attention for national and international data infrastructures have created a new data awareness amongst data producers. Suddenly they not only collect the data for their own prime task but have to make the data available to a wide range of users. A resent example in the hydrographic domain is the European EMODNET project [1] that will make bathymetric, chemical, physical and biological data of the European sea basin available to the general public. Frameworks underlying data infrastructures not only specify the why of these initiatives, they also specify the how by defining standard formats and protocols. Data producers now see themselves forced to comply with these frameworks and this requires a different view on the management of their data for more and different users.

2.3 The data management bottleneck

Today, software development in relation to modern multibeam echo sounders is primarily focussed on processing and analysis of (backscatter) data. These tools (e.g. Fledermaus from IVS-3D) provide a wealth of functionality that operate on the source data. Semi-automatic data cleaning and generation of surface hypotheses using the CUBE (Combined Uncertainty and Bathymetry Estimation) algorithm are just a few. CUBE is an error-model based, direct DTM generator that estimates the depth plus a confidence interval directly on each node point [2]. Most of these tools have a 3-D user interface (see Figure 1).

To face the ever increasing volumes of data, hydrographic organisations have invested in new tools to convert the raw sensor data into manageable work sets by applying some form of data reduction technique. Major characteristic of these tools is that they operate on survey files. For project based work (e.g. the processing of a single survey or the production of a paper chart) this is often sufficient but the limits of bathymetric file processing are coming closer. Not because (64 bit) operating systems can't handle very large files but the storage management of the data and the interaction with the stored data itself becomes a problem. To maintain acceptable read/ write access times for very large files advanced index and partitioning technology becomes inevitable. Bathymetric file management, even if it is just for single projects, therefore, will eventually evolve into data base management which is the traditional domain of the RDMS vendors.

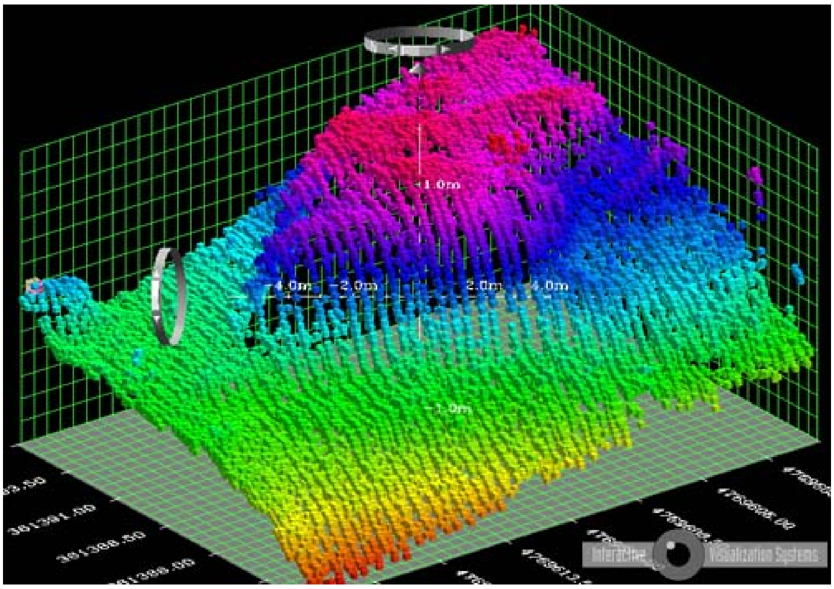


Figure 1. Advanced 3D editing using Fledermaus.

The huge volumes of (sensor) data do however create challenges for them too and the geo-spatial community must encourage data base vendors to continuously improve their systems.

2.4 Large digital terrain models

What is expected to become a problem for project based work is already pressing when it comes to building large high resolution digital terrain models (DTM) built from different raw point cloud sources. Creation of the DTM involves cleaning and processing of point cloud data but not automatically data thinning. Managing these models is therefore just as complex and challenging as managing the point clouds itself. As mentioned before the introduction of spatial infrastructures forces data owners to provide their collection of single data sets in a seamless and quality controlled way to the end-user. As long as these models have a relatively low resolution (e.g. 30 arc-second spacing) and up-to-dateness is not so much an issue, little problems exist. Examples of such data sets are the ETOPO1/GEBCO DTM (see Figure 2) and the Global Topography data set from Smith and Sandwell [3].

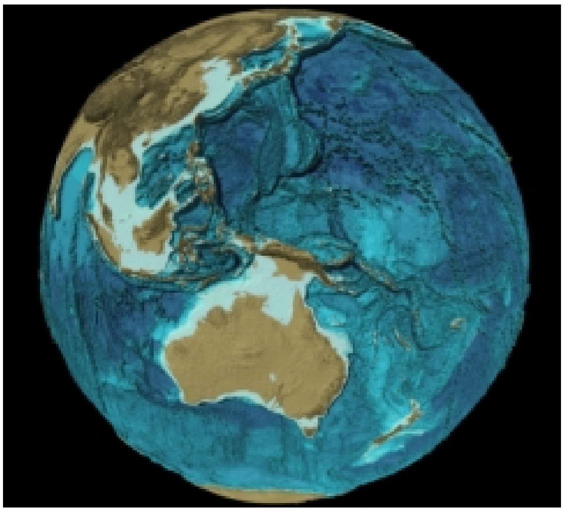


Figure 2. GEBCO(08) DTM.

Spatial data infrastructures however require more and better data. This is a key issue for SDIs and the reason they are being initiated and mandated (e.g. the European Inspire Directive [4]). Without better data decision makers can't make better decisions and they can't make them faster. Data sets have to be provided for a wide user community and this requires a careful planning of the management of the data. Provision of data in a spatial data infrastructure is an ongoing responsibility. It is not a one time project but end-users will expect almost direct access to all new data that comes available. This will be the challenge for many bathymetric data producers in the near future.

2.5 Morphological research

The challenge can probably be best explained by providing an example. Marine morphologists that study sediment transport in estuaries require many data sets representing the seabed topography at various moments in time. Various statistical sampling methods may be used to build the DTMs for this purpose but the main characteristic of these data sets is the fact that the data set represents a snapshot in time. If this would not be the case the risk exists that the same (moving) sediment is found multiple times in a single DTM.
The time factor is, therefore, the key issue and is dominant over the quality of the survey. In other words IHO S-44 – the survey (quality) standard of the International Hydrographic Organization [5] – quality is something a marine morphologist is less worried about if there are no alternative (better) data sets available. This implies that the models used by marine morphologists are different than the ones a marine cartographer is using for the production of a navigational chart. The expert in the latter example would not use lower quality data which is more recent if there is a slightly older higher quality data set available for the same area that provides better coverage.

To satisfy the needs of all users in a spatial data infrastructure flexibility will be the key issue. Given the fact that data volumes are increasing rapidly it will become a serious problem if an organisation collects various (end-user) requirements and tries to derive at agreed specifications for all these users.

2.6 Liability

Especially for National Hydrographic Offices responsible for producing navigational charts of their home water, liability is an important issue. Governments are responsible for the information they provide and can not easily wave that responsibility as most private (data) organisations do. In case of accidents caused by faults or omissions in navigational charts, the hydrographic office must provide evidence that the fault or omission can not be blamed on them. The only way to do that is to keep details records of their work and to operate a system that allows them to reproduce their knowledge of the situation before the accident occurred. Apart from archiving individual source documents and surveys, they also have to archive every change and process step associated with that change in the DTMs used for the production of the navigation charts.

Given the density of the datasets traditional backup procedures should be avoided as much as possible. So not just the various (end) user requirements call for efficient solutions but also the liability issue is heavily contributing to today's data management challenge.

3. Virtual Continuous Models

To overcome some of the problems associated with the management of large digital terrain models some solutions try to simplify the data as much as possible in order to reduce the size of the derived models and consequently increase the manageability. Although this may sound straightforward and workable, reality is different. Hydrographic data reduction by modelling the data sets into grids or TINs is lossy and the result is a compromise and does not always satisfy all users.

Concluding that storing many copies of high density models is not an option and building simplified models does not provide the solution either, calls for a solution that combines the advantages of both methods without incorporating the disadvantages. Virtual Continuous Models (VCM) as introduced by ATLIS combine the best of both worlds. A VCM defines the seabed topography based on rules defined by the user of the model. The rules are translated into standard SQL queries and describe how the individual surveys from the archive should de-conflicted and combined into a DTM without actually copying all the required data points into the DTM. An example of a simple rule could be: give the highest priority to the most recent survey and the lowest to the oldest. In order to increase the processing efficiency, each defined VCM is pre-processed by storing the outer limit or survey (concave) hull of those parts of the individual surveys that participate in the DTM. A concave hull follows the distribution of the point better than a convex hull. A concave hull can have so called dents bending inwards thus avoiding unnecessary "empty" spots inside the hull (see Figure 3 [6]). The hulls are stored unprojected as latitudes and longitudes to support potential world coverage of the VCM.

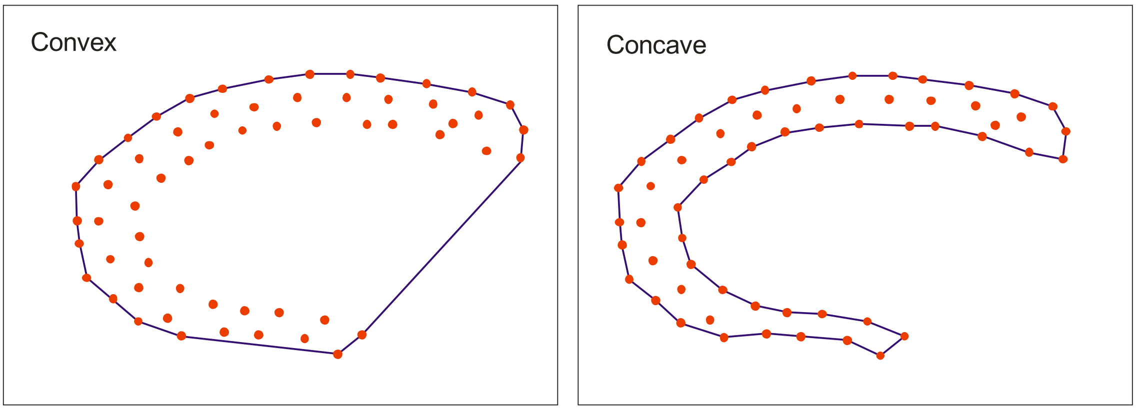


Figure 3. Convex versus concave hull.

The VCM definition consists of the rule used for de-conflicting and combining the individual modelled surveys and the set of concave hulls that together form the outer boundary of the DTM. This set can best be imagined as a source diagram that is often shown on paper navigation charts (see Figure 4).

The set of concave hulls for a given VCM can become very complex especially if many surveys have been conducted in the same area. The rule definition determines the stacking order of the individual surveys and also which data points from the source data participate in the result (e.g. shallowest, mean, CUBE hypotheses etc).

The number of VCM that can be defined using this mechanism is almost unlimited and the advantage is clear: the actual data points are only stored once.

3.1 Survey archive

The technology of VCM is based on the assumption that hydrographic organisations always archive the original source data. The idea is that if something goes wrong or turns out to be wrong in a later stage of the work process one can always revert back to the original data. Also the original data often contains much more than just depth information. Backscatter information can be geo-coded and used to determine bottom classification. This assumption is important as it does not in any way lead to more inefficiency in the solution as it requires no additional storage space. The source data is stored anyway but until now the archive was often difficult to access. Using VCM will change this. The only requirement for archiving the source data in order to make the data usable in the VCM concept is a common geodetic reference system. As a

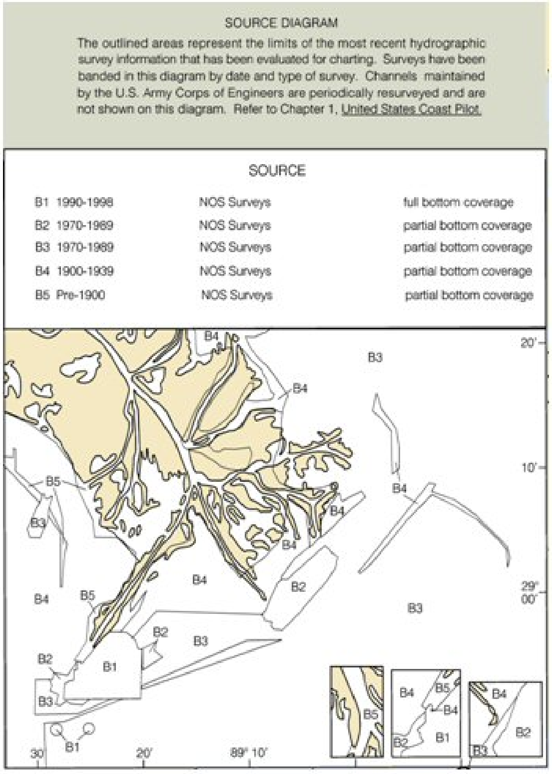


Figure 4. Example source diagram (courtesy NOAA).

DTM can cover the whole world, the system is based on a world coordinate system (lat/lon) and a single horizontal datum. Depths are negative and expressed in standard unit. Conversion of different coordinate systems and datums is performed as part of the archiving process. In theory this could be done on the fly but for performance reasons it is efficient to do this up-front. Vertical datum adjustment is also part of the modelling process.

The source data archive is catalogued by adding metadata to the individual surveys that allow easy access to the data based on international metadata standards (ISO 19115 and the OGC CSW standard).

3.2 Modeling process

Depending on the requirements of the organisation (and its customers) using VCM technology, the individual surveys can be modelled. Modelling may include to following steps:

- Adjustment (transformation) to a common geodetic model.
- Adjustment of vertical datum (based on separation models if available or fixed offsets).
- Data validation (e.g. using CUBE).
- Data thinning.
- Gap management (applying attributes to gaps in the data set describing the reason of the gap and the scale threshold for interpolation in products).
- Adding node level meta data (e.g. statistically derived quality information).

The result of the modelling process is called an Individual Model (IM) within the VCM concept. The IM acts as the building block of the VCM and has the same coverage as the original individual surveys. The IM is stored as a point cloud but is easier to manage as a result of the modelling process. This leads to a better overall performance of the VCM technology.

If required, the original (raw) data can be archived together with the modelled version of the individual survey.

3.3 Combining individual models

Different users require different rules for VCMs. This will result in different DTMs after applying the rules to the archive. Derived DTMs may even contain interpreted objects (e.g. interpolated depths) if the interpretation algorithm is added to the VCM definition. Figure 5 illustrates a survey archive consisting of 5 overlapping (conflicting) individual models.

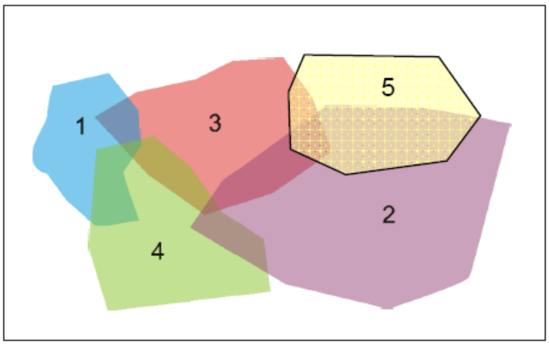


Figure 5. Archive of individual overlapping models.

After applying a rule set to the archive a VCM as shown in Figure 6 can be built.

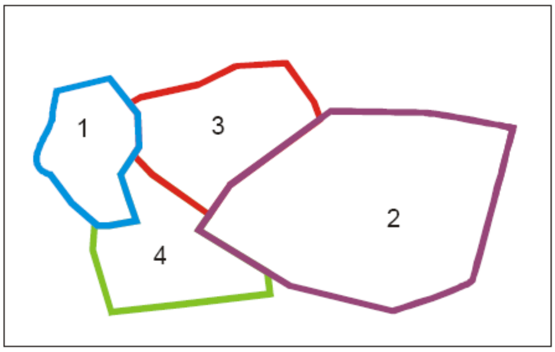


Figure 6. VCM.

The rule set used for the VCM in Figure 6 excludes IM 5 and give IM 4 the lowest priority. IM 1 and 2 have an identical priority and IM 3 has a higher priority than IM 4. Using a different rule set on the same archive could result in the VCM shown in Figure 7.

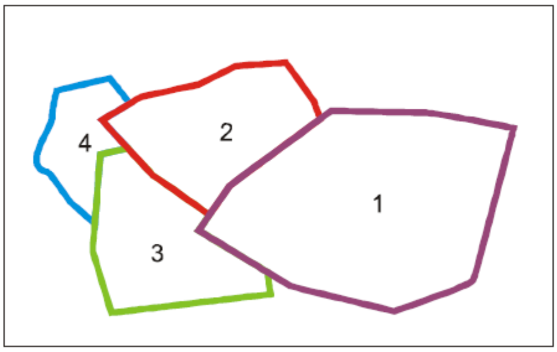


Figure 7. Alternative VCM.

The rules applied to the VCM definition can be very simple (e.g. most recent survey on top) or could be based on complex queries taking many meta data aspects into account. Although intersecting survey hulls is relatively straightforward in most GIS systems, care must be taken of the geodetic system used for the intersection. When using a non-projected coordinate system based on latitudes and longitudes, intersections must be based on geodetic lines if high accuracies are expected.

3.4 Performance

In order to make the VCM concept workable it is important to store the archive in a very efficient way. Fast data access to the data points is extremely important when the VCM is used in 2D or 3D visualisations or when output products (e.g. gridded bathymetry sets, contour lines or spot soundings) need to be generated.

ATLIS has implemented the VCM technology in SENS Bathymetry. The storage solution of SENS is based on Oracle spatial and utilises the latest in spatial technology that the Oracle 11g platform is offering today. Using the highly efficient spatial indexing technique combined with Oracle partitioning ensures a good performance of the VCM concept.

Performance is further enhanced by using the helical hyperspace (HHCode) concept for point data as originally designed for bathymetric applications by Mr. Herman Varma of the Canadian Hydrographic Service. The HHCode is a multidimensional construct based on the Riemannian Hypercube Structure that has an inherent z-order pattern also known as the Peano Curve [7]. The HHCode expresses dimensions in interleaved binary integer form to achieve high levels of compaction as well as rapid indexing and retrieval. Alternately stated, the HHCode is a recursive decomposition of an n-dimensional space. The entire planet Earth can be used as an example to illustrate the concept of the HHCode construction in 2D for a given spatial position (see Figure 8).

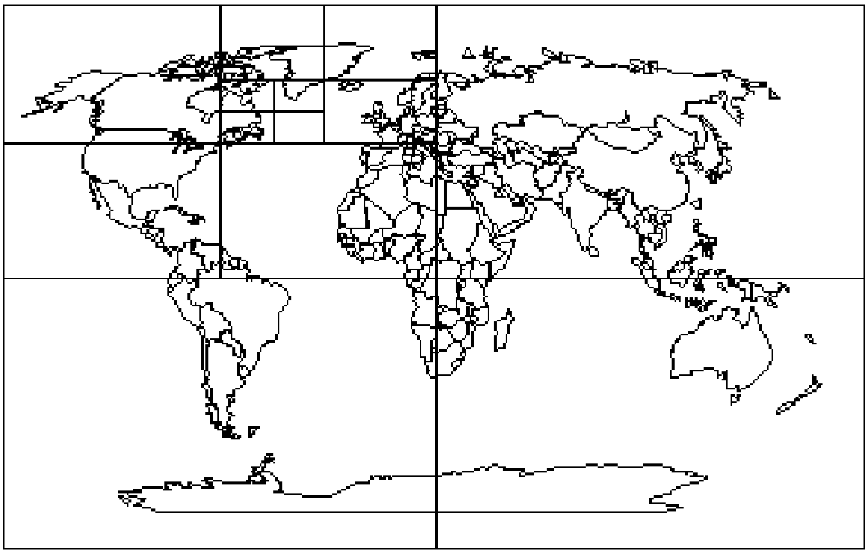


Figure 8. Recursive decomposition of space.

4. Future requirements

Even though the Oracle 11g platform provides enough power today to support the VCM concept, it is expected that the increase in data volumes of modern sensors system will continue to stress the limits of the storage capabilities. This requires on going research in further enhancing data access speed, indexing and partitioning technology. Especially the latest point cloud data structures introduced by Oracle are promising in that aspect.

5. Conclusions

Today's and future multi beam echo sounders and LiDAR systems generate a staggering amount of data. Although processing of huge individual survey files is still possible using advanced software packages like Fledermaus, managing the volumes of data in a seamless and unambiguous quality controlled way is becoming more and more a problem. Performance of today's system is no longer adequate given the data volumes and the increasing number of users that access data simultaneous. End users have different requirements and demand better data quicker. They are no longer satisfied with a fixed dataset built for general use. This requires the management and maintenance of multiple high resolution digital terrain models (DTMs). Given the huge volumes involved this will eventually lead to problems.

The Virtual Continuous Model technology, introduced by ATLIS provides a workable solution for this problem by no longer storing the data points redundant for each different DTM. The VCM only stores the definition of the required DTM by specifying the rules for combining and de-conflicting the individual surveys that are stored as individual models in the survey archive. Using the VCM concept gives a hydrographic organisation full control and maximum flexibility when it comes to managing bathymetric data. Participating in national and international spatial data infrastructures is no longer a problem but becomes an interesting show case for organisations using the VCM technology.

References

1. http://ec.europa.eu/maritimeaffairs/emodnet_en.html (May 2010).
2. Calder, B.R, "Automated Statistical Processing of Multibeam Echosounder Data", Submitted to Int. Hydro Review, January 2003.
3. Smith, W. H.F., and D.T. Sandwell, Global seafloor topography from satellite altimetry and ship depth soundings, Science, v. 277, p. 1957-1962, 26 Sept., 1997.
4. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).
5. IHO Standards for Hydrographic Surveys (5th edition), www.iho.org. (See also L.L. Dorst and C. Howlett, Evolutions in the IHO standards for hydrographic surveys, S44, Proceedings of the International Hydrographic Conference 'Evolutions in Hydro-graphy', Hydro06, Antwerp, Belgium, 6-9 November, 2006).
6. http://ubicomp.algoritmi.uminho.pt/local/concavehull.html (May 2010).
7. Varma et al., "A Data Structue for Spatio-Temporal Databases", International Hydrographic Review, Monaco, LXVII(1), Jan. 1990.

KONINKLIJKE NEDERLANDSE
AKADEMIE VAN WETENSCHAPPEN