# IMPROVING THE INTEGRATION OF DIGITAL SURFACE MODELS

**Jochen Schiewe**
University of Vechta, Germany
Institute for Environmental Sciences
jochen.schiewe@uni-vechta.de

**KEY WORDS:** DTM/DEM/DSM, Integration, Quality control, Formalization, Image matching, Object extraction

**ABSTRACT:**
Mutual benefits can be expected from an integrated processing of elevation and image data in order to improve the quality of automatically derived valued-added data. On this background new and modified algorithms will be described and empirically tested for the following closely related important tasks: The generation of Digital Terrain Models (DTMs) from Digital Surface Models (DSMs), the detection of buildings and wooded regions in the course of an object extraction, and the identification and revision of blunders within elevation models.

## 1 MOTIVATION

The user's acceptance of remotely sensed data mainly depends on the corresponding availability, quality and costs. The current status can be characterised in such a way that tremendous efforts have been made to deliver more data in shorter times by using new sensors and the benefits of the internet, while on the other hand a couple of important automatical data processing methods are neither reliable enough nor operational yet.

Additionally or complementary to traditional **data acquisition** methods electro-optical space sensors with improved spatial resolutions and along-track stereoscopic properties (e.g., with the Ikonos system), digital airborne scanners (e.g., the announced systems of Leica or Z/I Imaging), radar-interferometric sensors (e.g., the Shuttle Radar Topographic Mission) or laser scanners (e.g., the operational systems TopoSys or ALTM) became available in the last few years or will be on the market very soon. It is well known that the mentioned data sources are showing advantages and disadvantages which are in some cases complementary to the features of other sensors. For instance, laser scanning methods produce "blind data", i.e. no semantical or image information is associated with the elevation values - in contrast to optical sensors, whereas laser scanners are much better suited for capturing and processing heights in wooded or urban areas. Consequently, Ackermann (1999) forsees multi-sensor-systems that will represent a new development stage in the field of photogrammetry and remote sensing.

On the other hand, important tasks within the **data processing** chain like the automatical derivation of elevation models (e.g., by stereo matching) or the extraction of objects from imagery very often do not lead to satisfying and reliable results as obtained with human operators. Considering the above mentioned integration of multiple data sources new chances and challenges can be identified.

In this context, this paper wants to contribute to an improvement of the integration of image and elevation data – with the emphasis on the latter one, by presenting new and modified methodologies for key tasks. After a general integration concept is outlined (chapter 2), we will demonstrate and prove ideas and solutions for the generation of Digital Terrain Models (DTMs) from Digital Surface Models (DSMs), for the extraction of objects (with the emphasis on separating buildings from wooded areas), and for the detection and revision of height blunders (chapter 3). Chapter 4 will evaluate the achieved improvements and will give hints to future research and development needs.

## 2 CONCEPTUAL MODEL OF DSM INTEGRATION

Despite of the lack of actual multi-sensor-systems - with a very few exceptions for testing purposes (e.g., Hug, 1997), an integration of multi-sensor data from different sources and dates can already be performed. Following and generalizing the cited idea of Ackermann (1999), our **definition** of a multi-sensor integration including height information means

- either the fusion of DSMs of equal or different sources with equal or different spatial and/or temporal properties,
- or the fusion of DSMs with image data (especially in the visible and infrared electromagnetic spectrum).

In principle the integration of data can take place at all processing levels and it can be used for a variety of **purposes**, for example for

- automatically separating Digital Terrain Models (DTMs) from Digital Surface Models (DSMs);
- improving the automatical extraction of objects from imagery;
- detecting blunders resp. improving the reliability of elevation data;
- substituting missing height information (e.g., in regions of clouds or low sampling density);
- improving the extraction of morphological lines and points;
- generating virtual (static or dynamic) landscapes at different scales (Fritsch, 1999).

Of course, new **problems** are also arising with the integration of different data sets. For example, contradictory height or attribute information have to be handled. Hence, a key issue in DSM integration will be on blunder detection resp. error handling.

Following the idea of this conceptual model, it will be necessary to develop algorithms for the specific purposes in order to demonstrate or to disprove the desired profits from integrating elevation and image data. In the following we will concentrate on the fusion of DSMs with imagery by dealing with the first three of the mentioned topics which are actually closely linked to each other.

## 3 INTEGRATION TOPICS

In the following three closely related key topics with the integration of digital elevation and image data will be addressed: The derivation of the terrain surface from the DSM without additional information (section 3.2), which is for instance mandatory for the extraction of objects that stand out against their surrounding (like buildings or wooded areas; section 3.3). Vice versa these object information are also valuable for a correction of the approximated DTM as well as for the detection of blunders within the elevation model (section 3.4). Previously, section 3.1 describes the used data sets.

### 3.1 Test data

We will mainly rely on a data set covering parts of the City of Osnabrück (Germany) which was captured with the first version of the digital air-borne scanner HRSC-A (Scholten et.al., 1999). Because this scanner was originally designed for the use on the planet Mars, the spectral properties of the **imagery** especially in the red and infrared bands are not optimal for earth observation purposes (figure 1). In our case the spatial resolution of the scanner leads to a ground pixel size of 0.16 m for the panchromatic channel. The radiometric resolution of all channels amounts to 8 bit. In order to evaluate the intended object extraction a reference has been generated through visual on-screen digitizing.

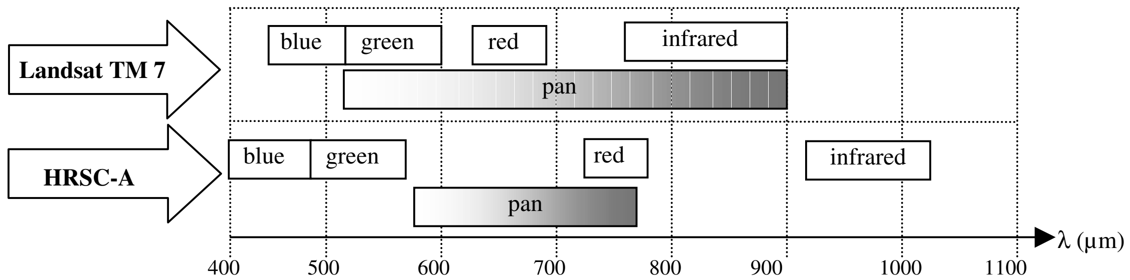



Figure 1. Design of spectral bands of HRSC-A scanner compared to those of Landsat TM.

Due to the nadir and the oblique looking (stereo) channels of the HRSC-A it is possible to acquire along-track stereoscopic **height data**. A gridded Digital Surface Model (DSM) with a horizontal spacing of 0.5 m has been obtained by automatical matching using ISTAR's SPOT3D software (Renouard and Lehmann, 1999). The DSM represents heights between 67 m and 100 m with an estimated accuracy of 0.1 m to 0.3 m (Scholten et.al., 1999).

In addition, also height data derived by automatical matching of imagery from the French satellite system SPOT covering the Westbank (Palestine) and Israel will be used.

### 3.2 DSM Normalization

In order to obtain heights relative to the terrain, a normalized Digital Surface Model (nDSM), i.e. the difference between the DSM and the Digital Terrain Model (DTM), has to be generated. In practise, neither the DTM nor any additional information (e.g., landuse data) are available, or these data are too expensive for an economic use, or they are not accurate or reliable enough, so that an approximated DTM has to be computed automatically.

**3.2.1 Discussion of previous work.** Apart from other methods for the normalization task (e.g., Shi and Shibasaki, 1996) the most widely cited and used procedure is the one which is proposed by Weidner and Förstner (1995): Based on mathematical greyscale morphology an *opening* consisting of an *erosion*, i.e. a minimum filtering on the original height values, and a *dilation*, i.e. a maximum filtering on this result, is performed on the DSM. The size of the moving window - also called the *structural element* ***W*** - has to be chosen in such a way that it is not entirely contained in the elements above the terrain which shall be cut off (i.e., buildings and wooded areas).

Obviously, the problem with this method is the proper choice of the size of ***W***: If it is set too small, the DTM becomes too high and the normalized DSM too low (see figure 2, left hand side). If this nDSM shall be used for object extraction (see section 3.3) valuable height information is lost (i.e., the number of omission errors is increasing). If on the other hand ***W*** is too large, the approximated DTM becomes too low and the nDSM too high, so that the "false alarm rate" (i.e., the number of commission errors) is increasing.

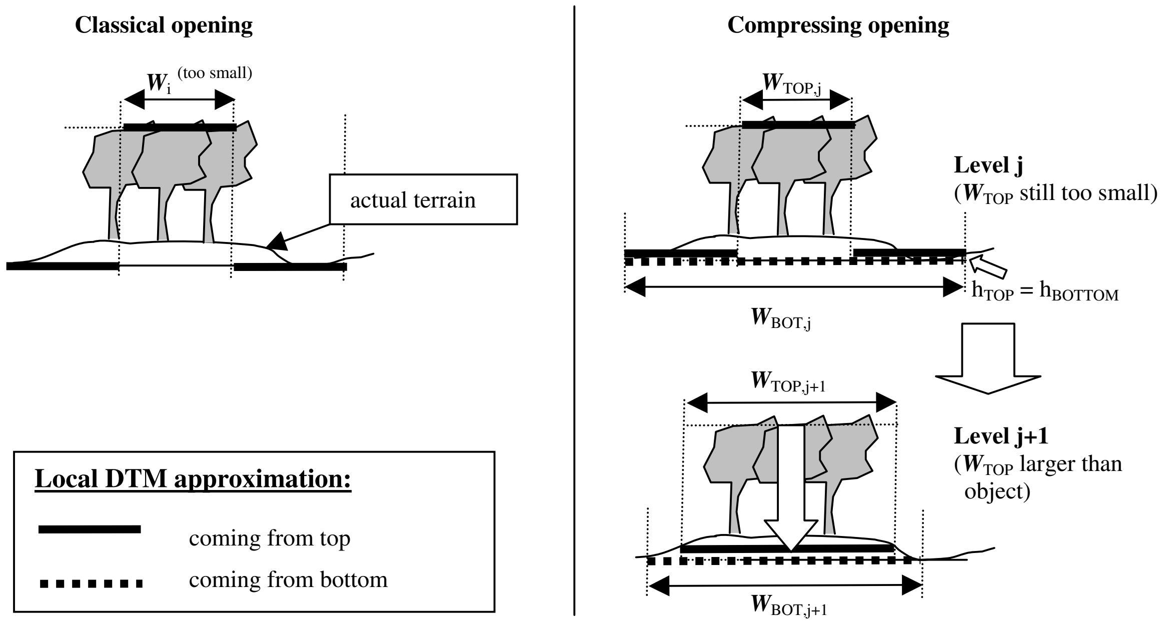



Figure 2. Principles of classical and compressing opening methods.

**3.2.2 Our approach.** In order to find a (sub-)optimal compromise between the mentioned extremes we propose the following method called *compressing opening*. This algorithm performs one opening from the top, i.e. starting with the DSM itself (resp. a structural element size $\boldsymbol{W}_{TOP}$ of 1), and one from the bottom, i.e. starting with the global minimum (resp. $\boldsymbol{W}_{BOTTOM}$ equals the image size). These openings are constantly repeated with a reduced $\boldsymbol{W}_{TOP}$ resp. an increased $\boldsymbol{W}_{BOTTOM}$ until both element sizes are equal. This proceeding leads to a monotonous decrease of the local minimum heights $h_{TOP}$ resp. an increase of $h_{BOTTOM}$ (i.e., the compression effect). If $h_{TOP}$ equals $h_{BOTTOM}$ – which has to be checked at every level, this is an indication that the terrain surface has been reached from both directions and hence the desired DTM approximation is obtained.

Figure 2 (right hand side) outlines the principle of the proposed method which in contrast to the classical opening indeed needs more computational time but delivers more reliable results. Figure 3 demonstrates a successful example showing deviations of the approximated DTM from the actual one of less than 1 m.

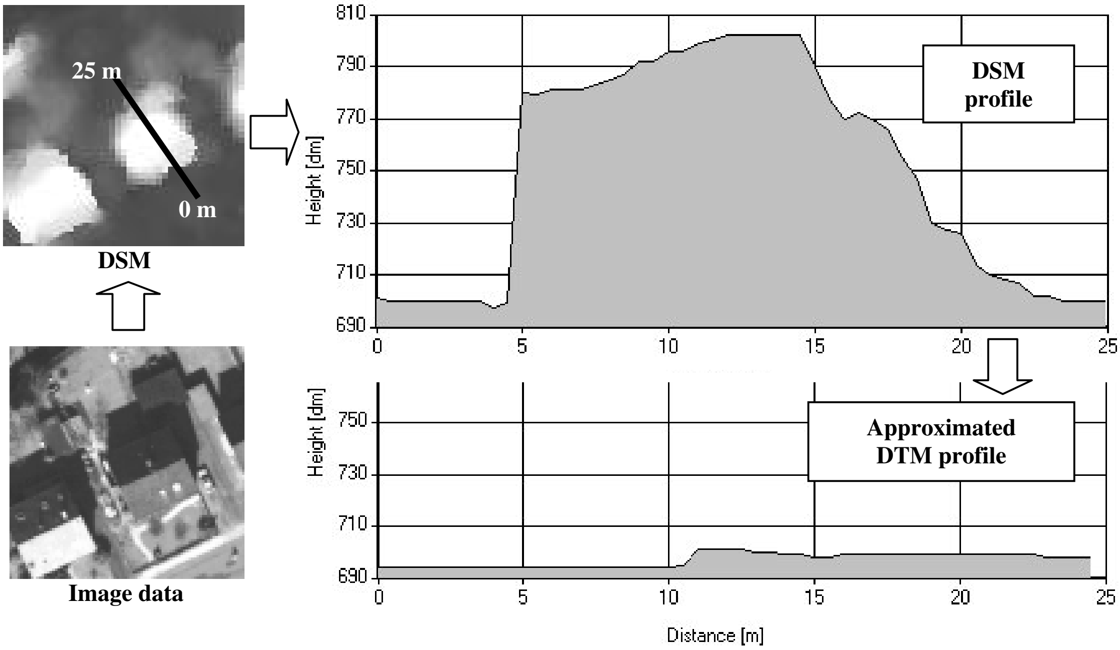



Figure 3. Example for DTM approximation through compressing opening.

### 3.3 Object extraction

Experiences from interactive stereoplotting have shown the importance of using elevation data in combination with spectral information for visually detecting and describing objects. The general idea is that buildings and wooded regions can be distinguished from roads and flat vegetated areas through a significant height difference (derived from a nDSM). The discrimination between classes of similar heights should be possible using spectral information (e.g., the Normalized Difference Vegetation Index, NDVI).

**3.3.1 Discussion of previous work.** Haala and Brenner (1999) propose a multispectral classification by introducing normalized height information as an additional channel. The disadvantage from our point of view is that the used unsupervised classification makes the process less transparency and not generally applicable. Hug (1997) uses a combination of laser altitude and reflectance data, thus reducing the spectral information content to a very narrow spectrum (in this case with $\lambda = 810$ nm). While other parameters like “elevation texture”, gradient variation or directional gradient distribution are of minor importance, the laser signal reflectance is found to be most useful for discriminating between buildings and forest regions. But in fact, the reported classification error rates (up to 23%) are by far not satisfying. Baltsavias et.al. (1995) propose a separation of buildings from trees by looking at the slope aspects which in the case of a single building show significant peaks at 90° apart from each other. - In practise, still a couple of problems - especially with the separation of buildings and wooded areas - occur with the described methods.

**3.3.2 Our approach.** Considering the advantages and drawbacks of existing approaches we will place the emphasis on the *detection* of buildings and wooded areas using a combination of various indicators derived from the normalized Digital Surface Model (nDSM) and from multispectral imagery. The decision for an object class (building, wooded area or others) is based on probabilities for the membership of a pixel to a certain object class with respect to every indicator which are then combined with a maximum a-posteriori estimation rather than relying only on single parameters or simple thresholding operations. In order to decrease the search space we rely on a two-stage concept that firstly differentiates between buildings *or* wooded regions against other objects, while the second stage tries to separate buildings from wooded areas.

In the first stage a “safe” **nDSM-altitude** threshold of 2.0 m is applied in order to differentiate between buildings (B) or wooded areas (W) against other objects. The probability of a pixel belonging to one of the desired classes can be defined for example by using the linear relationship

$$P(B \cup W \mid nDSM\_altitude) = \begin{cases} \dfrac{nDSM\_altitude - 2.0}{MAX(nDSM\_altitude) - 2.0} & if\ \ nDSM\_altitude > 2.0 \\ 0 & else \end{cases}$$

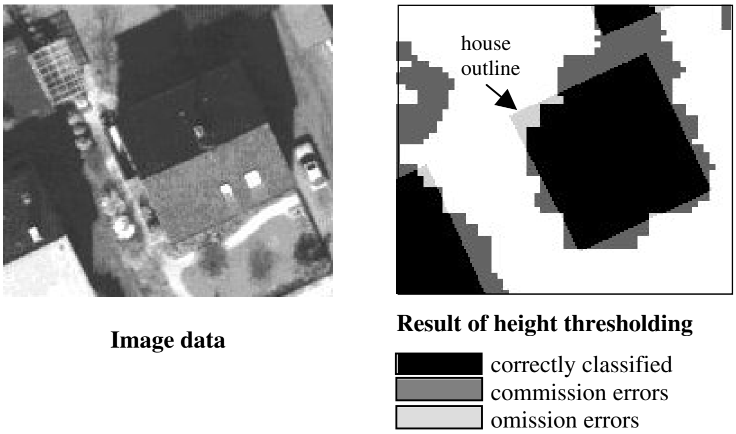



Figure 4. Classification using nDSM-altitude threshold.

Figure 4 demonstrates the result of this thresholding process: More than 99% of buildings or wooded areas are falling above this value, but the number of commission errors (i.e., other objects being also higher than this threshold) is rather large which is mainly due to errors within the DSM (especially in shadow regions, see also the DSM profile in figure 3).

Using this result the second stage tries to separate buildings from wooded areas (and potentially other objects) which fall into these candidate regions by means of various indicators that shall be shortly discussed in the following.

The hypothesis that the **density of slope gradients** is larger for wooded areas compared to buildings and other objects (see figure 5) can be formalized through the mean number of adjacent pixels with height gradients of more than 50° (which is larger than the presumable steepest roof slopes) within a local window (e.g., with 10•10 $m^2$ size). Our statistical analysis showed a good selectivity between classes as well as a good coincidence of the histograms with a normal distribution. Hence, the probabilities for the membership of a pixel to a certain object class can be determined by using the quantity $\Phi$ of the normal distribution, for instance for buildings (B):

$$P(B \mid gradient\_density) = \Phi(no\_of\_high\_gradient\_neighbours;\ \mu_B;\sigma_B)$$

The mean $\mu_B$ as well as the standard deviation $\sigma_B$ can be derived from training. After computing and comparing this measure for all object classes (building, wooded area, others), the maximum value indicates the most probable membership of the pixel to an object class. For our test site this method works very well with only few exceptions like for single trees or small lines of trees.

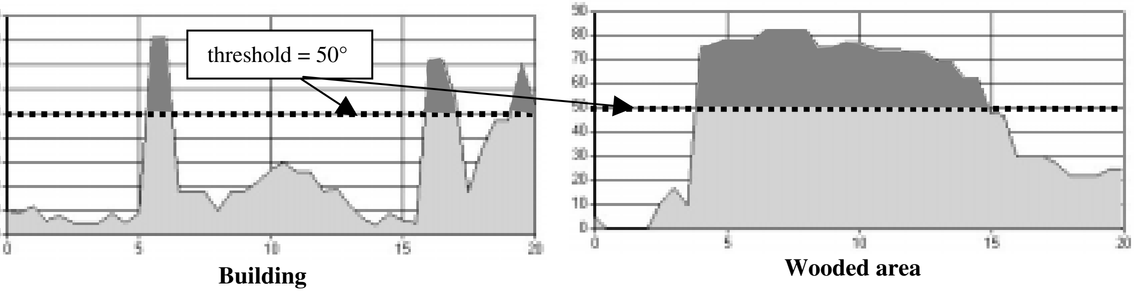


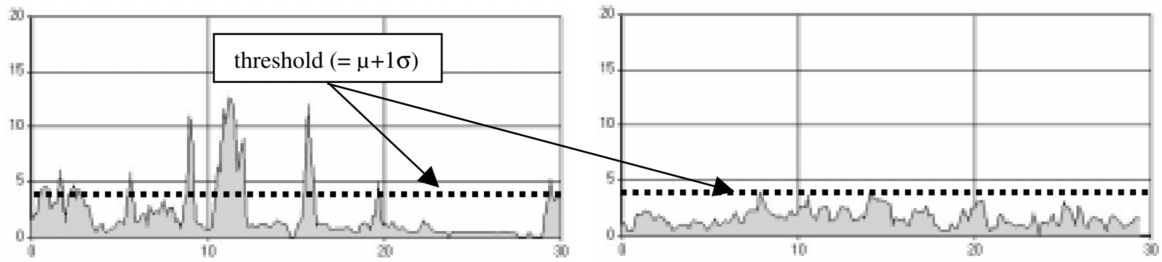

Figure 5. Slope gradient profiles for buildings and wooded area (gradients in degrees).

In principle, the **slope aspect** - for example proposed by Baltsavias et.al. (1995) - gives evidence of associated buildings if significant peaks 90° apart from each other can be observed, while the histograms of wooded areas do not show clear peaks. In the course of our tests this indicator is found to be not that useful because irregularly shaped buildings disturb the expected tendency and the number of gradient orientation values is too small for a reliable statistical evaluation.

With regard to spectral indicators, the **Normalized Difference Vegetation Index (NDVI)** is known to be a very good measure to differentiate between buildings and wooded areas. Unfortunately the above described design of spectral bands of the HRSC-A sensor (section 3.1) does not allow for a clear distinction of vegetated and non-vegetated areas. Also computing pseudo indices using other band combinations does not lead to better results. In principle, the probability of a pixel belonging to wooded areas could be described for instance with

$$P(W \mid NDVI) = \begin{cases} NDVI & if\ \ NDVI > 0.15 \\ 0 & else \end{cases}$$

Looking at the **spectral texture** the statistical analysis significantly proves that the highest values can be found with buildings because their breaklines are imaged much sharper than those of trees or other objects (figure 6). Applying the same principle as described with the slope gradient it is possible to separate very well between all object classes. Consequently, spectral edges could give evidence about the outline of buildings. But in practise, a couple of topological and geometrical processing steps have to be applied in advance, e.g. for the reduction of the extracted regions by shadow areas.




**Building** **Wooded area**

Figure 6. Spectral texture profiles of buildings resp. wooded area.

In order to find a final decision on the separation of buildings and wooded areas within this second stage all determined membership probabilities are compared with the goal to find a significant if not unique majority for one object class. This procedure also allows for the failure of one of the indicators (like the NDVI in our case) which is quite typical for real type applications and the use of different sensor configurations. It is also possible to integrate additional indicators (like shadow information) into the estimation.

With that the desired *detection* of buildings or wooded areas is finished and the following boundary *description* phase can be performed on basis of existing approaches like that of Sahar and Krupnik (1999), Shi et.al. (1997) or Fischer et.al. (1998) that partly also incorporate building hypotheses.

### 3.4 DSM control and generation

In order to improve the accuracy and - even more important - the reliability of automatically derived elevation models, an integration of other DSMs (e.g., through the merge of multiple correlations coming from a multiple image overlap) or of image and value-added data (in order to check the height behaviour of underlying object classes and their neighbourhood) can be used.

**3.4.1 Discussion of previous work.** Most studies are concentrating on the detection of blunders within single or multiple DSMs by means of geometrical parameters only. For instance, Lohmann and Koch (1999) define blunders by a certain deviation of a local plane. This method works well for point blunders, but worse for actual linear jumps like breaklines. Baltsavias et.al. (1995) present a merge of multiple correlations from a manifold image overlap on basis of the so-called “Figure of Merit” (FoM) which unfortunately gives unsatisfactory results.

**3.4.2 Our approach.** Concluding from the results concerning DSM control presented so far we highly recommend the integration of image data and image-derived information in order to try to check heights according to associated structures or objects rather than on relying on geometrical parameters only. In the following we will examine *spectrally homogeneous regions* which are critical areas for automatic matching algorithms. In this context, Krupnik (1998) proposes an interpolation of all interior heights. However, from our point of view homogenous regions should be treated in more detail, at least by grouping them into the following three categories (figure 7):

- Some regions (especially waters) demand not only for an interpolation but also for a constant height (i.e., $\Delta h = 0$);
- Some regions (especially shadows) do not allow any prediction about their heights – these have to be further processed or omitted;
- All other classes might be modelled through a plane or might show considerable height variations – depending on the underlying object type.

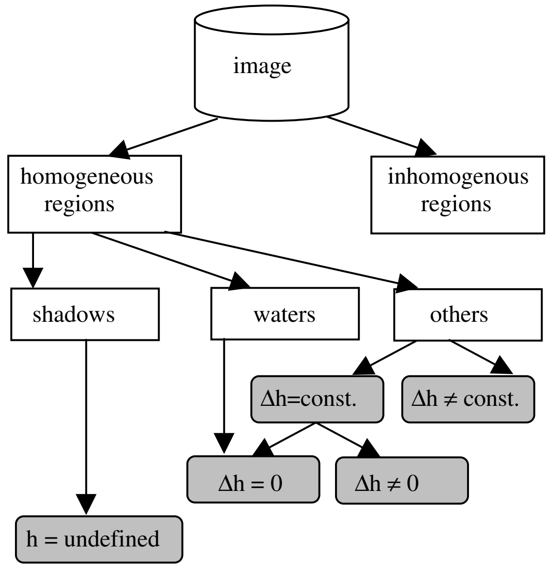



Figure 7. Height prediction for homogenous image regions.

To demonstrate a simple but typical example for the need of the proposed combined semantical and geometrical approach, figure 8 shows a height profile of an inland water within SPOT imagery (Schiewe and Isaac, 1999). The matching algorithm produces an undesired smooth transition from the surrounding to the water area, so that no constant height can be observed within the water body (in our examples height deviations come up to 40 m).

The identification of homogeneous areas can be done automatically using a variance filter. A fast elimination of this error would be to set all heights within the homogeneous region to the lowest value in this area. In order to improve the quality especially around the water boundary image preprocessing and topological post-processing steps can be applied (Schiewe, 1997).

A similar height prediction model can be set up for other critical image regions like *repetitive structures*.

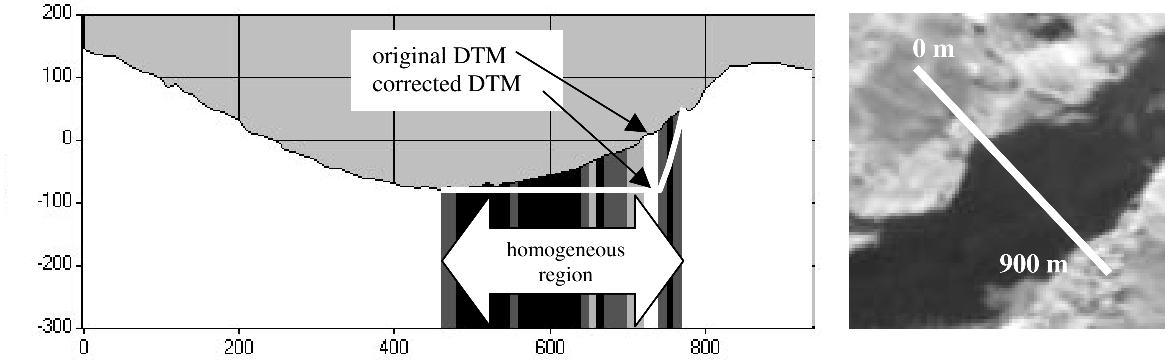



Figure 8. Detection of erroneous water regions within an elevation model (distance and heights in meters).

## 4 SUMMARY AND CONCLUSIONS

In order to achieve satisfying results by automatically processing remotely sensed imagery - for instance for the generation of reliable Digital Surface Models (DSMs) or the successful extraction of topographical objects - mutual benefits are expected from methods combining elevation and image data. We have presented and empirically proved modified and new procedures for different, but closely linked topics using digital airborne imagery:

- For the automatical determination of an approximated Digital Terrain Model (DTM) from a given DSM (resp. for computing a normalized DSM), we propose a new approach called *compressing opening* that in contrast to existing methods adjusts the size of local operator windows to the that of topographical objects which shall be eliminated.
- For detecting buildings and wooded areas in images we are using a combination of various elevation parameters (like normalized DSM altitude or density of slope gradients) and spectral indicators (like NDVI or spectral texture) by introducing probabilities for the membership of a pixel to a certain object class with respect to every indicator which are combined with a maximum a-posteriori estimation rather than relying only on single parameters or simple thresholding operations.
- For the detection of blunders within elevation models we recommend the integration of image or image-derived information in order to check the height behaviour accordingly to the corresponding structures or object classes and their neighbourhood (as a very simple example: water areas should show a constant height).

So far very promising results have been achieved by applying only rather simple image processing operations. Future implementations will also consider topological as well as advanced geometrical and semantical methodologies. Furthermore, it is obvious that the mentioned three processing topics are influencing each other - for instance, the object extraction phase gives evidence about objects that have to be eliminated for the generation of an approximated DTM which then can be re-computed, so that we will also implement an iterative and closed sequence of these topics which have been handled independently until now.

Finally it can be stated that the presented methods are significantly contributing to the initial goal of improving the integration of elevation and image data, and thus also to the envisaged use and success of upcoming multi-sensor-systems in the next few years.

## 5 ACKNOWLEDGEMENTS

The presented HRSC-A image data by courtesy of the City of Osnabrück, Germany (Fachdienst Geodaten). SPOT data by courtesy of partners within the project “GIS-based Models and GIS-Tools for Sustainable Transport Planning in Israel and Palestine” sponsored by the German Research Foundation (DFG).

## 6 REFERENCES

Ackermann, F., 1999. Airborne laser scanning – present status and future expectations. ISPRS Journal of Photogrammetry & Remote Sensing. Vol. 54, pp. 64-67.

Baltsavias, E., Mason, S. and Stallmann, D., 1995. Use of DTMs/DSMs and Orthoimages to Support Building Extraction. In: Gruen, A. et.al. (eds.): Automatic Extraction of Man-Made Objects from Aerial and Space Images. Birkhäuser Verlag, pp.199-210.

Fischer, A. et.al., 1998. Extracting Buildings from Aerial Images Using Hierarchical Aggregation in 2D and 3D. Computer Vision and Image Understanding, Vol. 72, No. 2, pp. 185-203.

Fritsch, D., 1999. Virtual cities and landscape models - what has photogrammetry to offer? In: Fritsch, D. and Spiller, R.H. (eds.): Photogrammetric Week ’99, pp. 3-14.

Haala, N. and Brenner, C., 1999. Extraction of buildings and trees in urban environments. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 54, pp. 130-137.

Hug, C., 1997. Extracting Artificial Surface Objects from Airborne Laser Scanner Data. In: Gruen, A. et.al. (eds.): Automatic Extraction of Man-Made Objects from Aerial and Space Images (II). Birkhäuser Verlag, pp. 203-212.

Krupnik, A., 1998. Automatic detection of erroneous areas in automatic surface reconstruction. In: International Archives of Photogrammetry and Remote Sensing, Vol. 32, Part 3, pp. 163-168.

Lohmann, P. and Koch, A., 1999. Quality Assessment of Laser-Scanner-Data. In: Proceedings of ISPRS-Workshop „Mapping from Space 1999“, Hannover, Germany, CD-ROM.

Renouard, L. and Lehmann, F., 1999. High Resolution Digital Surface Models and Orthoimages for Telecom Network Planning. In: Fritsch, D. and Spiller, R.H. (eds.): Photogrammetric Week ‘99, pp. 241-246.

Sahar, L. and Krupnik, A., 1999. Semiautomatic Extraction of Building Outlines from Large-Scale Aerial Images. Photogrammetric Engineering & Remote Sensing, Vol. 65, No. 4, pp. 459-465.

Schiewe, J., 1997. Suitability of MOMS-02 data for hierarchical map revision. In: Proceedings of ISPRS-Workshop "Sensors and Mapping from space 1997", Hannover, pp. 247-256.

Schiewe, J. and Isaac, J., 1999. Improving the Reliability of DEMs from Automatic Matching: Experiences from a case study. In: Proceedings of “2nd International Symposium on Operationalization of Remote Sensing”, Enschede (NL), CD-ROM.

Scholten, F. et.al., 1999. The high resolution stereo camera (HRSC) - digital 3D-image acquisition, photogrammetric processing and data evaluation. In: Proceedings of ISPRS-Workshop “Mapping from Space 1999”, Hannover, CD-ROM.

Shi, Z. and Shibasaki, R., 1996. Towards automated house detection from digital stereo imagery for GIS database revision. In: International Archives of Photogrammetry and Remote Sensing, Vol. 31, Part B4, pp. 780-785.

Shi, Z., Shibasaki, R. and Murai, S., 1997. Automated Building Extraction from Digital Stereo Imagery. In: Gruen, A. et.al. (eds.): Automatic Extraction of Man-Made Objects from Aerial and Space Images (II). Birkhäuser Verlag, pp. 119-128.

Weidner, U. and Förstner, W., 1995. Towards automatic building extraction from high resolution digital elevation models. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 50, No. 4, pp. 38-49.