Available online at www.sciencedirect.com
ScienceDirect
ISPRS Journal of Photogrammetry & Remote Sensing 61 (2006) 23–32
PHOTOGRAMMETRY & REMOTE SENSING
www.elsevier.com/locate/isprsjprs

Semantically correct 2.5D GIS data — The integration of a DTM and topographic vector data ☆

Andreas Koch, Christian Heipke *

Institut für Photogrammetrie und GeoInformation, Leibniz Universität Hannover, Nienburger Str. 1, 30167 Hannover, Germany

Received 7 July 2006; accepted 25 July 2006
Available online 11 September 2006

Abstract

The most commonly used topographic vector data, the reference data of national geographic information systems (GIS) are currently two-dimensional. The topography is modelled by different objects which are represented by single points, lines and areas with additional attributes containing information, for instance on the function and size of the object. In contrast, a digital terrain model (DTM) in most cases is a 2.5D representation of the earth's surface. The integration of the two data sets leads to an augmentation of the dimension of the topographic objects. However, due to inconsistencies between the data the integration process may lead to semantically incorrect results.

This paper presents a new approach for the integration of a DTM and 2D GIS vector data including the re-establishment of the semantic correctness of the integrated data set. The algorithm consists of two steps. In the first step the DTM and the topographic objects are integrated without considering the semantics of the objects. This geometric integration is based on a DTM TIN (triangular irregular network) computed using a constrained Delaunay triangulation. In the second step those objects which contain implicit height information are further utilized: object representations are formulated and the semantics of the objects are considered within an optimization process using equality and inequality constraints. The algorithm is based on an inequality constrained least squares adjustment formulated as the linear complementary problem (LCP). The algorithm results in an integrated semantically correct 2.5D GIS data set.

Results are presented using simulated and real data. Lakes represented by horizontal planes with increasing terrain outside the lake and roads which are composed of several tilted planes were investigated. The algorithm shows satisfying results: the constraints are fulfilled and the visualization of the integrated data set corresponds to the human view of the topography.

Keywords: GIS; DTM; Integration; Adjustment; Modelling

1. Introduction

1.1. Motivation

The most commonly used topographic vector data, the reference data of national geographic information systems (GIS) are currently two-dimensional. The

☆ Significantly enhanced version of a paper published in the proceedings of the International Symposium on Spatial Data Handling (SDH): Koch, A., Heipke, C., 2004. Semantically Correct 2.5D GIS Data — the Integration of a DTM and Topographic Vector Data. In: Fisher, P. (Ed.): Developments in Spatial Data Handling. Berlin, Springer, pp. 509–526.

* Corresponding author. Tel.: +49 511 762 2482; fax: +49 511 762 2483.
E-mail address: heipke@ipi.uni-hannover.de (C. Heipke).

0924-2716/$ - see front matter © 2006 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.isprsjprs.2006.07.005

topography is modelled by different objects which are represented by single points, lines and areas with additional attributes containing information, for instance on function and size of the object. In contrast, a DTM in most cases is a 2.5D representation of the earth's surface. The integration of the two data sets leads to an augmentation of the dimension of the topographic objects. However, due to inconsistencies between the data the integration process may lead to semantically incorrect results.

Inconsistencies may be caused by different object modelling and different surveying and production methods. For instance, vector data sets often contain roads modelled as lines or polylines. The attributes contain information on road width, road type etc. If the road is located on a slope, the corresponding part of the DTM often is not modelled correctly. The DTM points are distributed in a regular grid and break lines are often missing. When integrating these data sets, the slope perpendicular to the driving direction is identical to the slope of the DTM which usually does not correspond to the real slope of the road. In addition, the DTM and the GIS data are often produced independently. The DTM may have been generated using LIDAR (light detection and ranging) or aerial photogrammetry. Topographic vector data may be based on digitized topographic maps or orthophotos. These different methods may cause inconsistencies, too.

Many applications benefit from semantically correct integrated data sets. For instance, good visualizations of 3D models of the topography need correct data and are important e.g. for flood simulations and risk management. A semantically correct integrated data set can also be used to produce correct orthophotos in areas with non-modelled bridges within the DTM. Furthermore, a semantically correct integration may show discrepancies between the data and thus may allow to draw conclusions about the quality of the underlying DTM and the vector data.

1.2. Related work

The integration of a DTM and 2D GIS data is an issue that has been tackled for more than ten years. Weibel (1993), Fritsch and Pfannenstein (1992) established different forms of DTM integration: in case of *height attributing* each point of the 2D GIS data set contains an attribute "point height". By using *interfaces* it is possible to interact between the DTM programme and the GIS. Either the two systems are independent or DTM methods are introduced into the user interface of the GIS. The *total integration* or *full database integration* comprises a common data management within one database. The terrain data is often stored in the database in the form of a TIN whose vertices contain X, Y and Z coordinates. In general, however, the DTM is not merged with the data of the GIS. This merging process, i.e. the introduction of the 2D geometry into the TIN, has been investigated later by several authors (Lenk and Heipke, 2006; Lenk, 2001; Klötzer, 1997; Pilouk, 1996). The approaches differ in the sequence of introducing the 2D geometry, the amount of change of the terrain morphology and the number of vertices after the integration process. Stoter (2004) distinguishes between three different forms of integrated TINs: constrained TIN, conforming TIN and refined constrained TIN. They differ in the number of vertices and in the shape of the triangles near the integrated object geometries.

Among others, Lenk and Klötzer argue that the shape of the integrated TIN should be identical to the shape of the initial DTM TIN. Lenk developed an approach for the incremental insertion of object points and their connections into the initial DTM, represented as a TIN. The sequence of insertion is object point, object line, object point etc. The intersection points between the object line and the TIN edges (Steiner points) are considered as new points of the integrated data set (see Section 3).

Klötzer, on the other hand, first introduces all object points, then he carries out a new, preliminary triangulation. Subsequently, he determines the Steiner points and adds these points and the edges between the points to the data set. Since the Delaunay criterion is re-established in the preliminary triangulation, the shape of the integrated TIN may deviate somewhat from the one of the initial DTM. All methods have in common, that the mentioned inconsistencies between the data are neglected and thus may lead to semantically incorrect results.

Rousseaux and Bonin (2003) focus on the integration of 2D linear data such as roads, dikes and embankments into a DTM. The linear objects are transformed into 2.5D surfaces using attributes (road width and height) of the GIS database and the height information of the DTM. Slopes and regularization constraints are used to check semantic correctness of the objects. However, in case of incorrect results the correctness is not established. Only a new DTM is computed using the original DTM heights and the 2.5D objects of the GIS data.

2. Semantic correctness

2.1. Purely geometric integration

A DTM is composed of points with its coordinates X, Y, Z and an interpolation function to derive Z values at arbitrary positions X, Y. Mostly the DTM is a 2.5D representation of the topography, i.e. bridges, vertical walls and overhangs are not modelled correctly. Against this, the topographic vector data we consider are two-dimensional. The topography is modelled by different objects which are represented by single points, lines and areas.

Fig. 1 shows two examples of a purely geometric integration of a DTM and 2D topographic vector data (an integration without considering the semantics of the topographic objects). The left part shows some lakes, to the right a road network is shown. The integrated data set is represented by a TIN. The height values of the lakes do not have a constant height level. Several heights of the lakes near the bank are higher than the mean lake height. Also, the roads are not modelled correctly in the corresponding part of the DTM. The slopes of the cross sections are identical to the mean slope of the DTM. There are no breaklines at the left and right borders of the roads. Also, neighbouring triangles of the DTM TIN show rather different orientations. These representations of the lakes and the road do not correspond to the human view of the topography and thus represent a semantically incorrect result of the integration process.

2.2. Semantically correct integration

If we divide the topography into different topographic objects (roads, rivers, lakes, buildings, etc.), several objects have a direct relation to the third dimension. These objects contain implicit height information. For example, a lake can be described as a horizontal plane with increasing terrain at the bank outside the lake. To give another example, roads are usually non-horizontal objects. We certainly do not know the mathematical function representing the road height, but we know from experience and from road construction manuals that roads do not exceed maximum slope and curvature values in and across the driving direction.

Of course, all other objects are related to the third dimension, too. But it is difficult and often impossible to define general characteristics of their three-dimensional shape. For example, an agricultural field can be very hilly. It is not possible in general to define maximum slope and curvature values because these values vary from area to area.

The objects containing implicit height information, which needs to be used for a semantically correct integration, can be divided into three different classes (see Table 1). The first class contains objects which can be represented by a horizontal plane. The second class describes objects which are composed of several tilted and connected planes. The extent of these planes depends on the curvature of the terrain and the size of the object. Using the planes it is possible to adequately approximate the corresponding part of the original DTM. The last class shown in Table 1 describes objects which have a height relation to other objects. Bridges, tunnels and overpasses may contain a certain height relation to the terrain or water above or beneath. Whereas we deal with objects of the first two classes in the following, those of the third class are outside the scope of this paper.

To integrate a DTM and a 2D topographic GIS data set in a semantically correct sense, the implicit height information of the mentioned topographic objects has to be considered during the integration. This means, that the integrated data set must be consistent with our view of the topography. With a view to national reference GIS databases this means e.g., that all height values of points of the bounding polygon of a lake and all heights situated inside the bounding polygon must have the same height level, and the DTM points at the bank outside the lake must be higher than the lake height. Also, the slope and curvature of roads are bounded in and across the driving direction, the slope across the road can usually be

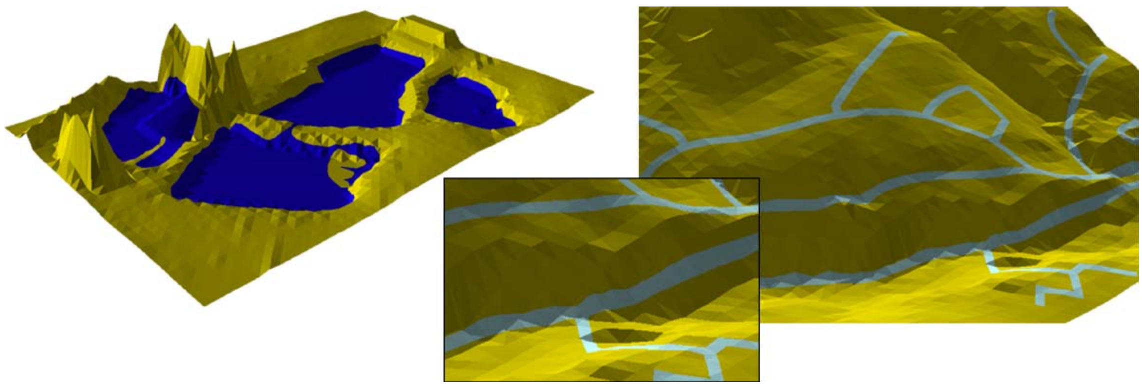

Fig. 1. The integration of a DTM and 2D topographic vector data without considering the semantics of the objects; left: lakes, right: road network.

Table 1
Some topographic objects and their representation in the corresponding part of the terrain

Object	Representation
Sports field	Horizontal plane
Race track	
Runway	
Dock	
Canal	
Lake, pool	
Road	Tilted connected planes
Path	
Railway, tramway	
River	
Bridge	Height relation
Tunnel	
Underpass, overpass	

neglected This means that points of a road cross section can be assumed to have constant height.

3. An algorithm for the semantically correct integration

The aim of the integration is to construct a consistent data set with respect to the underlying data model which complies with the semantics of the topographic objects in the sense described above. First, topographic objects, which are modelled by lines, are converted into elongated area objects, since they of course have a certain width in the landscape. This task is carried out through buffering, the buffer width is taken from the attribute width if available; otherwise a default value is used. After this pre-processing step the data sets are geometrically integrated without considering the semantics of the topographic objects (Section 3.1). The integration is based on a DTM TIN computed using a constrained Delaunay triangulation (Lee and Lin, 1986). The object borders are split by introducing the intersection points between the DTM TIN and the object geometries (Steiner points) and the split object geometries are introduced as edges of the triangulation. The result is an irregular triangular network — an integrated DTM TIN. Then, certain constraints are formulated and are taken care of in an optimization process (Section 3.2). In this way, the topographic objects of the integrated data set are made to fulfill predefined conditions related to their semantics. The constraints are expressed in terms of mathematical equations and inequality constraints. The algorithm results in improved height values, which lead to a semantically correct, integrated 2.5D topographic data set.

A basic assumption of our approach is that the general terrain morphology as reflected in the DTM is correct and has to be preserved also in the neighbourhood of objects carrying implicit height information. Therefore, any height changes resulting from the integration must be as small as possible, and the neighbourhood of the objects has to be taken into account in order to guarantee a smooth transition between changed and non-changed areas. A second assumption is that inconsistencies between DTM and topographic objects stem from inaccurate DTM heights and not from planimetric errors of the topographic objects. We have introduced the latter assumption in order to separate the two potential causes of error. It is clear that if this assumption is violated, the results produced by our method may not be of much use.

3.1. Geometric data integration

There are several approaches for the integration of a DTM and 2D topographic GIS data based on a triangulation. The advantage of Lenk's approach (Lenk and Heipke, 2006; Lenk, 2001) compared to others is that the surface shape represented by the integrated TIN is identical to the one of the initial DTM TIN. The disadvantage is that the approach results in a large amount of Steiner points which lead to additional observation equations and/or inequality constraints (see Section 3.2). Since we require the changes of the original heights to be as small as possible (see above), we have chosen to use Lenk's method, improving it with respect to the shape of the triangles of the integrated TIN. For the sakes of completeness the algorithm is shortly described in the following: Starting from a DTM TIN the 2D object points and their connections are inserted: The location of the first starting polygon point is determined, and then the intersection points between the triangle sides and the object polygon are calculated and introduced into the polygon. As a result, new polygon sides are created, which lie exactly inside the triangle. Similar to Lenk any redundant points are now discarded. The polygon sides are considered as edges to be preserved, and both, the left and the right part of the object polygon inside the triangle, are re-triangulated using a polygon triangulation. In this step the shape of the triangles is recursively optimised with respect to equal angles: for all possible solutions the smallest triangle angle is determined, and the solution with the largest such angle is selected. After processing this triangle the neighbouring triangle in direction of the object polygon is considered. The first intersection point is known from the calculations before: the second one is now computed and the integrated TIN is derived in the same way. This process is repeated for all 2D objects. The

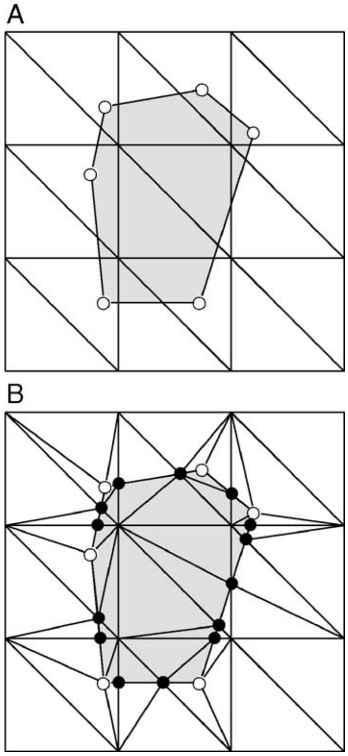

Fig. 2. Integration of a DTM and a topographic object “lake”, A) original DTM TIN and object “lake”, B) integrated data set.

result of this geometric integration is unique because of the exhaustive search in the polygon triangulation.

Fig. 2 shows an example of the integration of a DTM and an object “lake” of a 2D GIS data set. The original points of the bounding polygon of the lake are shown in white. After the integration, the intersection points between the DTM TIN and the object polygon are new points of the integrated data set (depicted in black).

Another example is given in Fig. 3. The shown road is modelled by lines, which is buffered using the road width (Fig. 3A). First, the intersection points between the road axis and the DTM TIN are introduced as new points of the road object (see black points). This is done because every triangle may have a different inclination and the road axis should be best fitted to the terrain represented by the DTM TIN. The left and right side of the buffered road, which contain as many points as the road axis including the new points, are then introduced using the described variant of Lenk’s algorithm (Fig. 3B).

3.2. Optimization process

As mentioned, there are topographic objects of the 2D GIS data which contain implicit height information. Within the integrated data set these objects have to fulfill certain constraints which can be expressed in terms of mathematical equations and inequality constraints. To fulfill these constraints and thus to achieve semantic correctness, the heights of the DTM are adapted. As mentioned before, the horizontal coordinates of the polygons of the topographic objects are introduced as error-free in our current approach.

The heights of the topographic objects and the neighbouring heights outside the objects are considered as unknowns and are estimated within an optimization process which is based on a least squares adjustment. The heights of the corresponding part of the DTM are introduced as direct observations for the unknown heights at the same planimetric position. Equality constraints are introduced using pseudo observations, and the adherence to the constraints is controlled via weights for the pseudo observations. Furthermore, inequality constraints are formulated to described additional constraints. The optimization process is solved using the linear complementary problem (LCP) (Lawson and Hanson, 1995; Fritsch, 1985; Schaffrin, 1981).

3.2.1. Basic observation equations

The heights of the DTM which correspond to the topographic objects of the 2D GIS data (heights inside

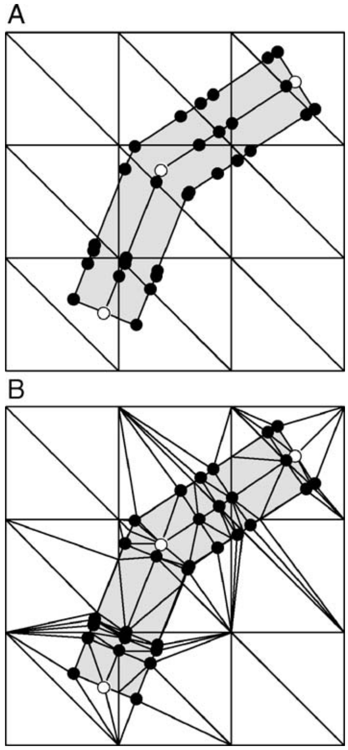

Fig. 3. Integration of a DTM and a topographic object “road”, A) original DTM TIN and object “road”, B) intersection between DTM TIN and object.

the objects, heights of points of the bordering polygon, heights of neighbouring points outside the objects) are introduced as:

$$\hat{v}_i = \hat{Z}_i - Z_i \tag{1}$$

The height Z_i refers to the original height of the DTM, the value $\hat{Z}_i$ denotes the unknown height which has to be estimated, $\hat{v}_i$ is the residual of the observation i. In case of a point of the bounding polygon Z_i has to be interpolated using the height information of the DTM.

In order to be able to preserve the slope of an edge connecting two neighbouring points P_j and P_k of the DTM TIN where one is part of the polygon describing the object, and the other one is a neighbouring point outside the object (and thus to constrain changes to the general shape of the integrated DTM TIN) additional equations are formulated:

$$\hat{v}_{jk} = \hat{Z}_j - \hat{Z}_k - (Z_j - Z_k) \tag{2}$$

3.2.2. Equality and inequality constraints

Each class of object representation (see Table 1) has its own constraints which can be expressed in terms of mathematical equality and inequality constraints. These constraints are derived next for the two types of planes.

3.2.2.1. Horizontal plane. Heights of objects which represent a horizontal plane must be identical everywhere. This means that points P_l with height Z_l and planimetric coordinates X_l, Y_l situated inside the object boundary (see Fig. 4A, grey points) must all have the same value $\hat{Z}_{\mathrm{HP}}$ which has to be estimated in the optimization process. These height values lead to the following observation equation:

$$\hat{v}_l = \hat{Z}_{\mathrm{HP}} - Z_l \tag{3}$$

Since the points P_m of the bounding polygon of the topographic objects with planimetric coordinates X_m, Y_m do not have a value Z_m, heights are derived from the DTM. We linearly interpolate Z_m, from the three neighbouring values Z_u, Z_v, Z_w. Again, the height difference between the unknown object height and the interpolated height is used to formulate a pseudo observation (see Fig. 4A, black points):

$$\hat{v}_m = \hat{Z}_{\mathrm{HP}} - Z_m(X_m, Y_m, Z_u, Z_v, Z_w) \tag{4}$$

The neighbouring terrain of the horizontal plane is considered using the basic observation equations, Eqs.

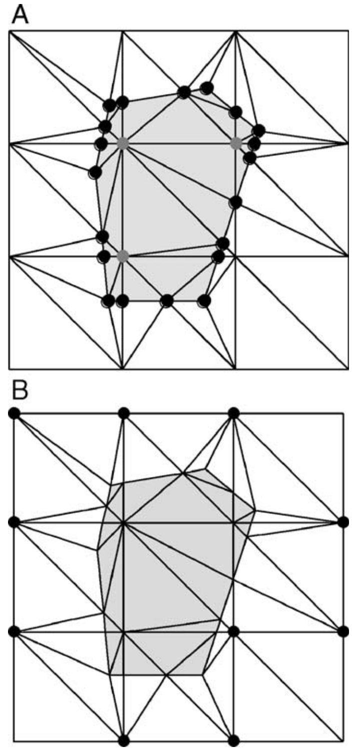

Fig. 4. Equality and inequality constraints of a horizontal plane, topographic object "lake", A) points inside the lake and points of the waterline, B) points of the neighbouring terrain.

(1) and (2) (see Section 3.2.1). If the object represents a lake it is necessary to use a further constraint which represents the relation between the lake in terms of a horizontal plane and the bank of the lake whose height values $\hat{Z}_i$ have to be higher than the height level of the lake $\hat{Z}_{\mathrm{HP}}$:

$$0 > \hat{Z}_{\mathrm{HP}} - \hat{Z}_i \tag{5}$$

In Fig. 4B the points $\hat{Z}_i$ of Eq. (5) which are points of the neighbouring terrain are shown in black.

3.2.2.2. Tilted planes. The objects treated in this paper which can be composed of several tilted planes are roads. The example in Fig. 5 shows a road which is modelled by a centre line and then buffered using the attribute "road width". In longitudinal direction roads are not allowed to exceed a predefined maximum slope value $s_{\max}$:

$$s_{\max} \geq \left| \frac{\hat{Z}_n - \hat{Z}_o}{D_{no}} \right| \tag{6}$$

Here, $\hat{Z}_n$ and $\hat{Z}_o$ are the unknown height values of successive points P_n and P_o of the road axis (Fig. 5A).

D_{no} is the horizontal distance between these points. In addition, the difference between two successive slope values which is comparable to the curvature of the object is restricted to the maximum value $ds_{\max}$:

$$ds_{\max} \geq \left| \frac{\hat{Z}_n - \hat{Z}_o}{D_{no}} - \frac{\hat{Z}_o - \hat{Z}_p}{D_{op}} \right| \tag{7}$$

P_n, P_o and P_p are successive points of the road axis, D_{no} and D_{op} are the corresponding horizontal distances. The values $s_{\max}$ and $ds_{\max}$ are based on road construction manuals or on experience.

Assuming a horizontal road profile in the direction perpendicular to the axis the height values of corresponding points must be identical:

$$\hat{v}_{nq} = \hat{Z}_n - \hat{Z}_q \tag{8}$$

The values $\hat{Z}_n$ and $\hat{Z}_q$ represent the heights of the axis and the left or the right side of the buffered object, respectively (Fig. 5B). These constraints are introduced for all cross sections whose centre point results from the intersection between the DTM TIN and the original object line. In Fig. 5B these cross sections are p_1 and p_3.

Those cross sections whose centre points are original points of the object axis are not used to form this kind of constraint because in the original points the road may show a change in horizontal direction and slope (cross section p_2 in Fig. 5B). Consequently, the cross section may be inclined.

Finally, the points of any two neighbouring cross sections and the points in between are constrained to lie in a plane:

$$0 + \hat{v}_r = \hat{a}_0 + \hat{a}_1 X_r + \hat{a}_2 Y_r - \hat{Z}_r \tag{9}$$

In Fig. 5B the points of the neighbouring profiles p_1 and p_2 as well as the points in between form the input to Eq. (9). These points P_r have to lie on the plane described by the unknown coefficients $\hat{a}_0$, $\hat{a}_1$, $\hat{a}_2$. X_r, Y_r are the planimetric coordinates of P_r, $\hat{Z}_r$ is the height of P_r which has to be estimated.

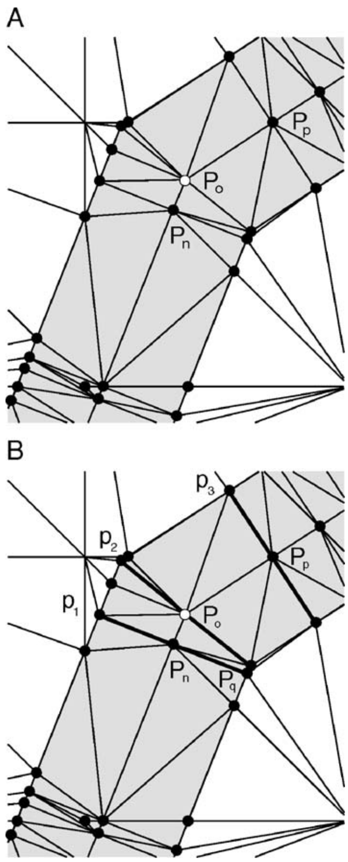

Fig. 5. Equation and inequation constraints, A) maximum slope and maximum slope difference, B) horizontal profile and points belonging to a plane.

3.2.3. Inequality constrained least squares adjustment

The basic observation equations (Section 3.2.1) and the equation and inequality constraints (Section 3.2.2) are introduced in an optimization process which is based on an inequality constrained least squares adjustment. The stochastic model of the observations (basic observations and equation constraints) consists of the covariance matrix of the observations. Assuming that the observations are independent of each other, this matrix has a diagonal structure and contains only the variances of the observations. The algorithm is formulated as the linear complementary problem (LCP) which is solved using the Lemke algorithm (Lemke, 1968). For more details see Koch (2003), the LCP is explained in detail in Lawson and Hanson (1995), Fritsch (1985) and Schaffrin (1981).

In the LCP, the inequality constraints are treated as hard constraints, which are automatically fulfilled. To fulfill also the equation constraints the related observations have to be considered to have a high accuracy, and the corresponding diagonal elements of the weight matrix (the inverse of the covariance matrix) have to be large. The constraints formulated for the integration are then fulfilled, if the residuals of the equation constraints are negligible. Thus, the semantic correctness of the resulting integrated data set depends on the choice of the weights of the equation constraints compared to those for the basic observation equations, where the latter

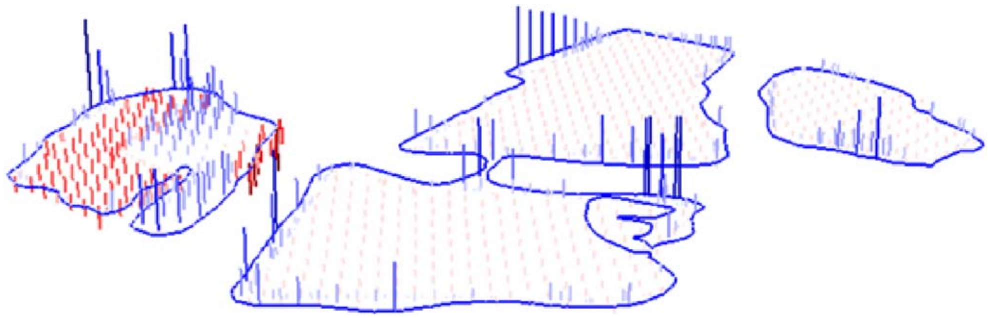

Fig. 6. Height differences between the original heights of the DTM and the estimated heights of the optimization process (vertical exaggeration factor: 30), object: lake.

should be based on the quality of the input height information.

4. Results

In this chapter we present results of some experiments conducted to study the behaviour of the developed integration method for different kinds of input data. The aims are to show (1) how the results depend on the weights of the basic observations and the equation constraints using simulated data, and (2) the behaviour of the algorithm using real data. We have investigated two different object classes — lakes, which are represented by horizontal planes, and roads which are composed of several connected tilted planes.

4.1. Simulated data

For the simulation we first created an artificial landscape with a 10 m DTM and integrated a lake with 221 height values and three roads to it in a semantically correct manner. The heights near the lake were chosen to be a little larger than the water heights. The distance between the lake and the roads was large enough, so that the objects did not influence each other. We then added white noise with a standard deviation of 0.5 m to all heights. This value was also introduced as the standard deviation for the basic observation equations, Eqs. (1) and (2).

In the case of horizontal planes, theory dictates that the result must always be semantically correct, because regardless of the weights for the Eqs. (1), (2), (3) and (4) one can compute a mean value for the plane from the estimated heights, and this mean value represents the lake height. Because of the introduced noise some of the heights of the neighbouring points (the points near the lake) became smaller than the mean value of the original heights of the object points (the heights of the lake border and those inside the lake). As a consequence, some of the inequality constraints Eq. (5) are not fulfilled. This situation changes during the optimization in two ways: the heights on the banks become larger, and the height of the lake is slightly reduced compared to the mean value of the input values. In this way a semantically correct result is achieved, as was to be expected. Furthermore, the empirical standard deviation of the heights inside the lake are nearly identical to the accuracy of the DTM because the residuals directly represent the differences between the estimated lake height and the original DTM height values.

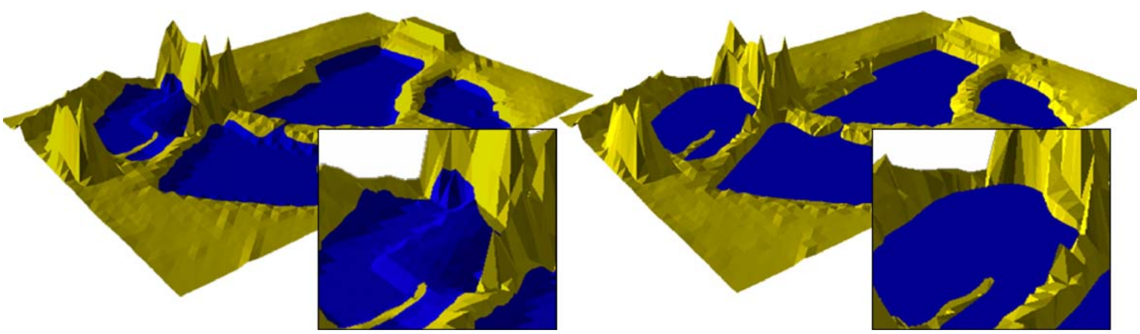

Fig. 7. Results of the integration, left: without considering the semantics of the topographic objects, right: semantically correct integration.

The road data set consists of three objects and a DTM with 315 grid points. In order to obtain a semantically correct result, the weights for the equation constraints Eqs. (8) and (9) were set to high values, the standard deviations of the basic observation equations, Eqs. (1) and (2), result from the height accuracy of the input DTM. In this case the residuals of the equation constraints obviously also become small, but the residuals of the road border points, the points inside the roads, and the original DTM points are larger. Up to a standard deviation of about 0.2 m for the equation constraints Eqs. (8) and (9) their theoretical standard deviation turned out to be smaller than 0.05 m, which for most applications can be considered to be small enough to call the result semantically correct.

The results show, that in case of lake the results always are semantically correct, but they depend on the quality of the input data. In case of a road the ratio between the weights of the equation constraints and the basic observation equations is of crucial importance in order to obtain a semantically correct result.

4.2. Real data

The real data were taken from the German ATKIS.[1] Three different lakes from an area in Lower Saxony were used, the dataset covers an area of 450 times 650 m^2. The lake borders consist of 294 planimetric polygon points, the DTM contains 1.961 grid points with additional 1.047 points representing structure elements (break lines). A semantically correct integration was carried out by using standard deviations of 0.5 m for the basic observation equations, Eqs. (1) and (2), and 0.1 m for the equality constraints, Eqs. (3) and (4). The number of basic observations and equation constraints was 2.754; 533 parameters had to be estimated and the number of inequality constraints was 530. The results show, that all constraints were fulfilled after carrying out the optimization. The differences between the estimated lake heights and the initial mean height values of the object points are very small. The first mean height of the first lake is reduced by 2 mm, that of the second one by 4 mm. The third lake is 3.7 cm lower than the original mean height value. This larger value is caused by a higher number of heights at the bank which initially did not fulfill the inequality constraint (Eq. (5)).

Fig. 6 shows the residuals after the optimization process. The blue vectors correspond to height values which are lower than the original heights. Red coloured vectors refer to heights which became higher. The figure shows that most of the heights inside the lakes became higher. Most of the points which became lower are situated at the border of the lakes. Against it, a big part of the differences of the left lake became lower, too. Here, the corresponding part of the DTM seems to be erroneous. The maximum differences between the original heights and the estimated heights are -1.84 m and $+0.88$ m, respectively. Fig. 7 shows the result of the semantically correct integration (right side) with respect to the results without considering the semantics of the lakes (left side, identical to Fig. 1). The semantically correct integrated data set shows that all constraints are fulfilled; the height values inside the lake and at the water line have the same level, the terrain outside the lake rises.

5. Conclusions and outlook

This paper presents an approach for the integration of a DTM and 2D topographic GIS data including the re-establishment of the semantic correctness of the integrated data set. The algorithm is based on a Delaunay triangulation and a least squares adjustment taken inequality constraints into account. Investigations were carried out using simulated and real data sets. The objects used in our investigations are lakes represented by a horizontal plane with increasing terrain outside the lake, and roads which can be composed of several tilted planes. The results are satisfying. In case of a lake the results always are semantically correct, the actual result (the lake height) depends on the quality of the input data. For roads the semantic correctness depends on the relation between the weights of the equation constraints and the basic observations. In the investigated example, this choice turned out not to be very critical, however.

Height blunders are not modelled and can therefore lead to non-realistic, yet in the sense of the paper semantically correct results. Thus, blunders have to be detected and corrected prior to the overall adjustment. Furthermore, the planimetric coordinates of the topographic objects were introduced as error-free. This may cause an erroneous height level of the topographic objects. Another issue which shows up implicitly in our research is the conversion of spatial dimension of topographic vector data such as road data from 1D to 2D. While we use the buffer method to do so in our work, the conversion of spatial dimension is a much more general problem which

[1] ATKIS stands for Authoritative Topographic Cartographic Information system and represents the German national geospatial reference database. The DLMBasis (basic digital landscape model) contains the highest resolution an, approximately equivalent to a topographic map 1:25,000, the DTM DGM is a hybrid dataset containing regularly distributed points with a grid size of 12.5 m and additional geomorphological information and an accuracy of approximately 0.5 m in Lower Saxony (AdV, 1989).

needs considerable attention when extending our approach to other object classes. In addition, planimetric coordinates of structure elements have to be explicitly considered in the adjustment algorithm. Otherwise, the resulting surface may deviate from the one represented by the initial DTM. We currently work on these issues, and also develop an extension which incorporates planimetric shifts for inaccurate 2D vector data in the adjustment procedure.

Acknowledgement

This research was generously supported by the State Surveying Authority of Lower Saxony *Landesvermessung und Geobasisinformation Niedersachsen* (LGN). We also express our gratitude to LGN for providing the data.

References

AdV, 1989. ATKIS Gesamtdokumentation, erstellt vn der Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland. Landesvermessungsamt Nordrhein-Westfalen, Bonn.

Fritsch, D., 1985. Some additional information on the capacity of the linear complementary algorithm. In: Grafarend, E., Sanso, F. (Eds.), Optimization and Design of Geodetic Networks. Springer, Berlin, pp. 169–184.

Fritsch, D., Pfannenstein, A., 1992. Conceptual models for efficient DTM integration into GIS. Proceedings EGIS'92, Third European Conference and Exhibition on Geographical Information Systems, pp. 701–710. Munich.

Klötzer, F., 1997. Integration von triangulierten digitalen Geländemodellen und Landkarten. Diploma thesis, Institute of Informatics, Rheinische Friedrich–Wilhelms—Universität Bonn, unpublished.

Koch, A., 2003. Semantically correct integration of a digital terrain model and a 2D topographic vector data set. Proceedings of ISPRS Workshop on Challenges in Geospatial Analysis, Integration and Visualization II, Stuttgart, 7 p., on CD-ROM.

Lawson, C.L., Hanson, R.J., 1995. Solving Least Squares Problems. Society for Industrial and Applied Mathematics, Philadelphia ISBN 0-89871-356-0. 337 pp.

Lee, D.-T., Lin, A.-K., 1986. Generalized Delaunay triangulations for planer graphs. Discrete and Computational Geometry 1, 201–217.

Lemke, C.E., 1968. On complementary pivot theory. In: Dantzig, G.B., Veinott, A.F. (Eds.), Mathematics in the Decision Sciences, Part 1, pp. 95–114.

Lenk, U., 2001. 2.5D-GIS und Geobasisdaten- Integration von Höheninformation und Digitalen Situationsmodellen. PhD Thesis, Wiss. Arb. Fachr. Verm. Hannover University, No. 244 and German Geodetic Commission (DGK) at the Bavarian Academy of Sciences, Series C, No. 546.

Lenk, U., Heipke, C., 2006. The radial topology algorithm — a new approach for deriving 2.5D GIS data models. Geoinformatica 10 (4), to appear.

Pilouk, M., 1996. Integrated Modelling for 3D GIS. PhD Thesis. ITC Publication Series No. 40, Enschede.

Rousseaux, F., Bonin, O., 2003. Towards a coherent integration of 2D linear data into a DTM. Proceedings of the 21st International Cartographic Conference (ICC), pp. 1936–1942. Durban.

Schaffrin, B., 1981. Ausgleichung mit Bedingungs-Ungleichungen. AVN 6, 227–238.

Stoter, J., 2004. 3D cadastre. Netherlands Geodetic Commission. Publications on Geodesy, vol. 57. 327 pp.

Weibel, R., 1993. On the integration of digital terrain and surface modeling into geographic information systems. Proceedings 11th International Symposium on Computer Assisted Cartography (AUTOCARTO 11), Minneapolis, Minnesota, pp. 257–266.