Fusion of LIDAR Data and Aerial Imagery for Automatic Reconstruction of Building Surfaces

Martin Huber[1], Wolfgang Schickler[2], Stefan Hinz[3], Albert Baumgartner[3]

***Abstract* – An approach for building reconstruction based on fused information extracted from different data sources, namely LIDAR data (LIght Detection And Ranging) and aerial imagery, is proposed. The building reconstruction is performed within the scope of a general surface estimation process. This surface estimation aims at generating a DTM including buildings and vegetation removed. The buildings are reconstructed by applying polyhedral models. Thus a large variety of building types can be described with the only limitation, that the roof surface consists of planes.**

***Index terms* – Building Reconstruction, LIDAR, Aerial Imagery, Fusion**

I. Introduction

Three dimensional city models are important for a large number of applications. More and more users of Geographic Information Systems (GIS) require precise 3D city models. Applications can be found, e.g., in the field of city planning, for simulations, for microclimate investigations, or in telecommunications for transmitter placement. This strong demand for 3D city models led to increasing efforts in developing 3D building reconstruction methods and, to minimize the costs for the 3D city modeling, research is going towards an automated reconstruction. First good results are reported from semi-automated and partially from fully automated algorithms [1].

A couple of years ago the main input data for the production of 3D city models were aerial images, terrestrial images, map data, and data derived from classical surveying [2]. Since resent years LIDAR (LIght Detection And Ranging) has become a very attractive alternative for the acquisition of 3D information. This technique directly provides a high density of 3D points. But due to the discrete and irregular distribution of these points, LIDAR is not able to capture features like straight breaklines or ridges directly. Aerial images, however, store information about the electro-magnetic radiance of the complete scene in almost continous form. Therefore they support the localization of breaklines and linear or spatial objects. The motivation for this work was to combine the complementary properties of aerial imagery and LIDAR data.

1) German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Oberpfaffenhofen, D – 82234 Wessling.
E-mail: Martin.Huber@dlr.de

2) Sanborn L.L.C., 1935 Jamboree Drive, Co 80920, USA.
E-mail: wschickler@sanmap.com

3) Chair for Photogrammetry and Remote Sensing, Technische Universität München, Arcisstraße 21, D – 80290 Munich.
E-mail: {Stefan.Hinz}|{Albert.Baumgartner}@bv.tum.de

The building model, which is the foundation of the whole reconstruction process, will be specified in Sect. 2. Section 3 explains the strategy as well as the basic steps and the characteristic elements of the approach. An evaluation of the results and a short outlook conclude the paper (Sect. 4).

II. Building Model

For the task of building reconstruction, we first of all have to define the geometric building model. This model defines the set of real world buildings, which can be dealt with. Brenner [3] comes to the conclusion, that the combined parametric and the polyhedral models were the most suitable ones for automatic or semiautomatic building reconstruction. Both provide a good compromise between the reconstruction effort and the number of supplied real world building types. Combined parametric models imply geometric and topologic characteristics. They are based on constructive solid geometry (CSG). A large number of different building types can be modeled, especially buildings with different eave heights. However, unusual building and roof types cause problems. With polyhedral models, most building roofs can be described. Such models can capture a complex roof topology and geometric conditions can be introduced explicitly by equations. The only limitation to a polyhedral model is that its roof surface must consist of planes (e.g., no dome-shaped surfaces are supplied).

In our approach a polyhedral building model is applied in order to be able to describe a large variety of building types. The model is split up into different, semantically meaningful substructures, which are the basis for building reconstruction. Following substructures and relations between substructures are defined (see also Fig. 1):

Each building consists of a *building roof* and *walls*. A building roof is defined by a minimum/maximum size and has to lie on the top of walls. The building walls are a set of *wall planes* having a minimum slope. The transition between the roof and the walls is the *roof boundary*. The building roof consists of single *roof planes*, which have a minimum size and include a minimum angle with the wall planes. Each roof plane has a *roof plane surface* marked off with a closed *roof plane boundary* (polygon consisting of single *roof plane outlines*). The lines where two adjacent roof planes touch are called *common roof plane outlines*. All other lines which are only part of one roof plane are called *single roof plane outlines* and are therefore part of the roof boundary.

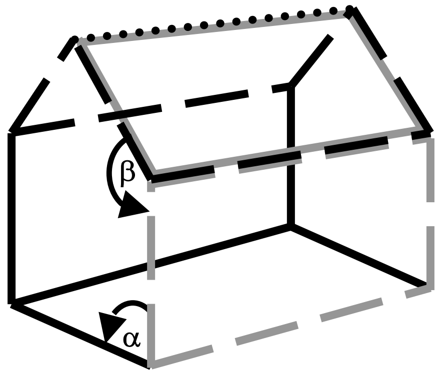


Figure 1: Building model.

Roof boundary (dashed black line) separating *roof* from *wall*. *Wall plane* (dashed gray line) having a certain slope α. *Roof plane boundary* (solid gray line) encloses a *roof plane surface* having a minimum size and minimum angle (β) with the wall plane. A *common roof plane outline* (dotted line) connects two roof planes.

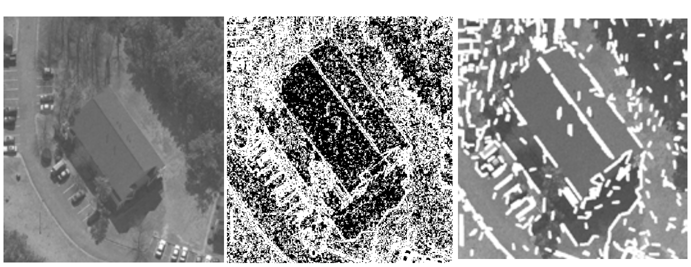
III. Automatic Reconstruction of Building Surfaces

The building reconstruction is embedded in a surface estimation process [4] and realized in two iterations. The first estimation focuses on the coarse detection and the second aims at the exact reconstruction of buildings.

A. Building Detection

Our goal in the building detection step is to obtain the approximate position and shape of buildings. Therefore, we estimate a preliminary surface, where vegetation and outliers in the LIDAR data are removed and only a DTM (Digital Terrain Model) with building-blobs is left. In order to receive this preliminary surface, information gained from the fusion of features extracted from aerial imagery as well as from LIDAR data is put into the surface estimation process.

Feature Extraction from image and LIDAR data: From aerial imagery, two types of features are extracted using conventional image processing algorithms: homogeneous regions (”*Image-Blobs*”) and gray values edges (”*Image-Edges*”), see also Fig. 2. From the raw LIDAR data, three different features are extracted providing hypotheses about vegetation, buildings, and breaklines: “*LIDAR-TreeBlobs”, LIDAR-RoofBlobs*”, and ”*LIDAR-Edges*”.



(a) (b) (c)

Figure 2: a) Aerial Image of test building. b) Result of blob extraction (dark). c) Result of edge extraction (white).

The first LIDAR feature (“*LIDAR-TreeBlobs*”) is based on the fact, that each laser pulse provides a first and a last return. Height differences between the two returns can be easily computed and enable us to draw conclusions regarding the measured surface. Solid surfaces have only a slight height difference whereas vegetation shows a larger difference between the two measured returns. Based on this information areas with vegetation can be detected. The second LIDAR feature (”*LIDAR-RoofBlobs*”, see Fig. 3) provides information about possible roof planes. Its extraction is split up in two processes, a region growing and a geometric reasoning. In the region growing process, the triangulation of raw LIDAR points is examined for planar surfaces, because the building model we want to reconstruct consists exactly of such planar surfaces. In the following pairs of triangles sharing one side were examined for plainness by estimating a plane on basis of the raw LIDAR points involved. Thus every point has a certain distance to the estimated plane. If the largest distance is still smaller than a certain threshold, then the triangles are merged. In the next step all adjacent triangles, which share one side with the already merged triangles, are examined one after another. As long as an adjacent triangle exists, which fulfills the criterion of plainness, the plane is updated and keeps on growing. After the region growing has finished the resulting planes are inspected according to their slope to classify them into candidates for roof planes and wall planes within a geometric reasoning process. We start this geometric reasoning by searching for steep planes. Although not each steep plane is automatically part of a building – they can also appear in areas with vegetation or where the terrain itself is steep – steep planes are a good starting point to find triangles belonging to the roof region. Hence, all neighboring planes of the steep ones are now examined with regard to certain geometric criteria (e.g. minimum height difference, angle between the normal vectors, size of a building; cf. Fig. 1). After merging planes fulfilling these criteria, the result of the geometric reasoning are closed building roof blobs consisting of all triangles classified as belonging to the roof. Finally, the third feature (”*LIDAR-Edges*”, see Fig. 4) is also based on last returns and provides areas with height discontinuities. This feature indicates the 2D position of 3D breaklines. Therefore

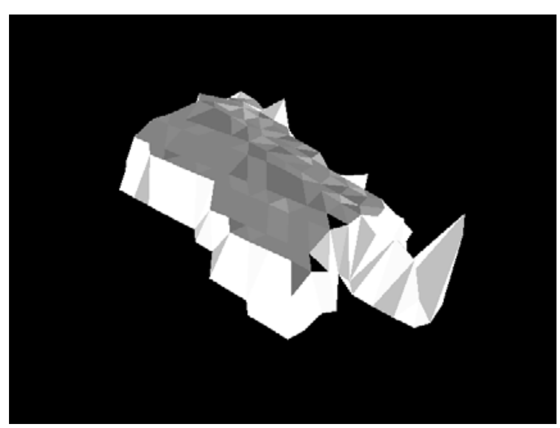
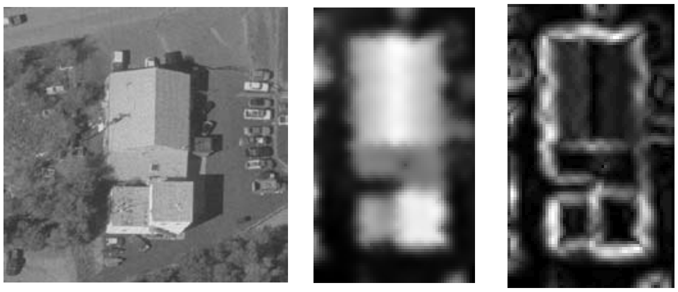


Figure 3: *LIDAR-RoofBlob* and wall triangles.



(a) (b) (c)

Figure 4: Building example 2. a) Aerial image. b) Range data represented as gray value image. c) LIDAR-Edges (gradient magnitude image).

a range image based on the raw LIDAR data is created by interpolation. By applying image processing algorithms the gradient magnitude image of the range image can be calculated containing information about height discontinuities.

Information Aggregation: In the next step previously extracted features are fused in order to get hypotheses for the position and shape of the buildings. The features LIDAR-TreeBlobs, LIDAR-RoofBlobs, and Image-Blobs are projected into a common co-ordinate frame and a label image is generated. Since each label indicates different characteristics of the terrain (e.g. vegetation, buildings), this image will be called "Surface Class Image" and will be used later to control parameters and weights involved in the surface estimation process. In the next step Image-Edges are projected into the ground system by intersecting them with the raw LIDAR point triangulation (see Fig. 5). Each Image-Edge is compared with the LIDAR-Edges, in order to get a more precise statement whether the projected Image-Edge could be a real 3D breakline (e.g. building breakline) or whether it is only a edge between different gray values in the aerial image (e.g. shadow on the ground, street markings). Therefore a projected Image-Edge is compared with the gradient magnitude image of the LIDAR-Edges. In case of a high gradient magnitude it gets a higher weight in the following surface estimation process.

Preliminary Surface Estimation: The goal of the Preliminary Surface Estimation (PSE) is to derive a DTM with buildings included and vegetation removed. It is based on three different input components:

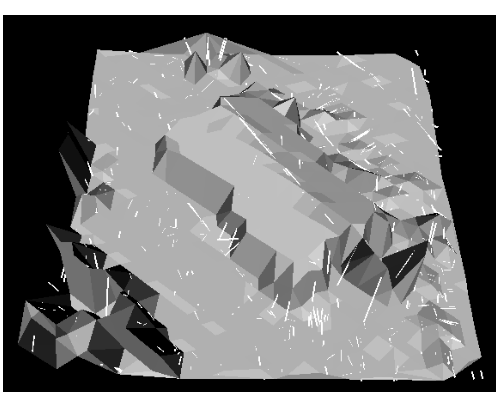


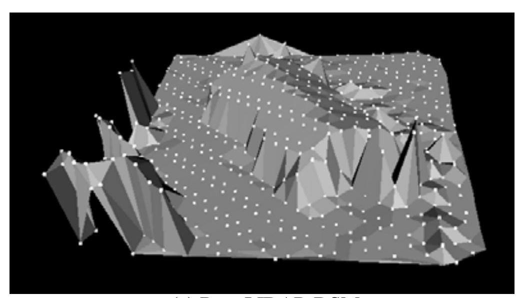
Figure 5: DSM with intersected "Image-Edges".

The raw LIDAR points (last returns) represent the genuine information. They are the basis for the whole estimation procedure. The Surface Class Image contains information about which parameter set is applied to which region. Parameter sets are defined for LIDAR-TreeBlobs, LIDAR-RoofBlobs, and Image-Blobs as well as for every combination of the three. The extracted Image-Edges are also added to the estimation process. Their weight depends on the surface class region they fall in. The input parameters for the surface estimation process define the functional and stochastic model of the estimation and have a strong influence on the resulting surface. A parameter set consists of the standard deviation of the individual LIDAR points, standard deviation of the surface curvature constraint, standard deviation of the surface slope constraint, and parameters of the one-sided robust estimation function [4,5].

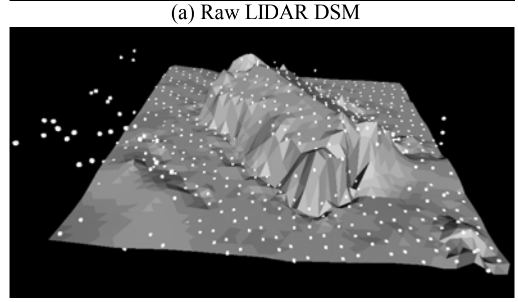
Although only the approximate building position and shape is detected in this first iteration, it is important since "noise" like vegetation and outliers are removed (see Fig. 6). Thus the preliminary surface can serve as input information for the more accurate building reconstruction.

B. Building Reconstruction

The goal in the building reconstruction step is to estimate a DTM including buildings with sharp outlines and well-defined roof structures. In contrast to the first estimation, which relied on the raw LIDAR-data, the second estimation is based on the preliminary estimated surface. Hence it is guaranteed, that vegetation and outliers are already removed. Thus the position and the structure of the building roof can be found more precisely. The processing chain for building reconstruction is quite similar as for building detection and consists of three phases: Building Feature Extraction, Information Aggregation, and Final Surface Estimation.



(a) Raw LIDAR DSM



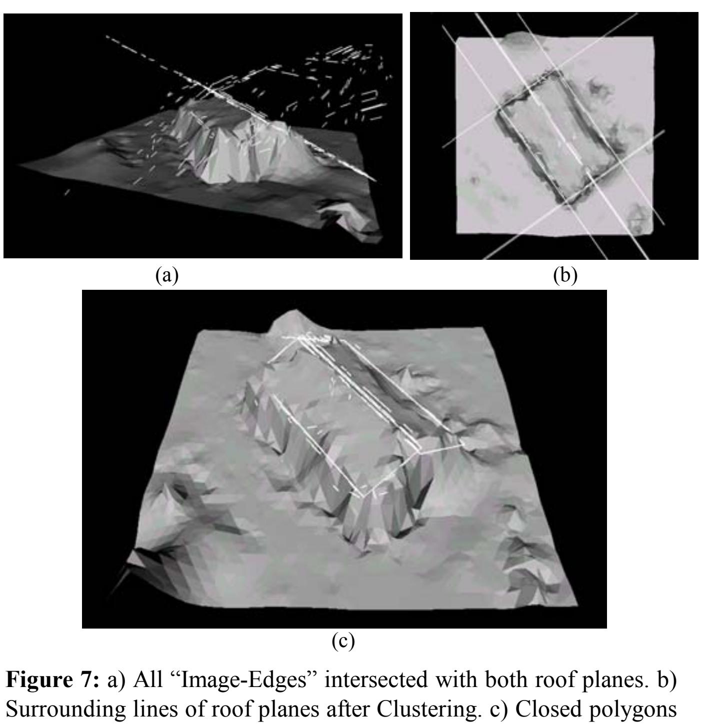
(b) Preliminary estimated surface

Figure 6: Superimposistion of LIDAR points and surfaces

Building Feature Extraction: Instead of classifying the whole roof region like in the LIDAR-RoofBlobs extraction for building detection (cf. Sect. 3 A), we now want to extract each roof plane separately. To this end, we extract planes and edges with the same algorithms as before. The difference is, that the extraction is now based on the preliminary estimated surface, and that the involved parameters are tightened. To get a clear differentiation to the previous feature we now call it "*PSE-RoofPlanes*" instead of LIDAR-RoofBlobs and "*PSE-Edges*" instead of LIDAR-Edges.

It is not necessary, however, to extract the other features again, because they are based on the raw input data. The LIDAR-TreeBlobs have already been extracted from the raw LIDAR data and the Image-Blobs as well as the Image-Edges have been extracted from the aerial imagery, respectively.

Information Aggregation: The main goal in this section is to find the 3D boundaries of the individual roof planes. Each roof plane is assumed to be plain and to have an outline polygon consisting of straight lines. Through the fusion of the PSE-RoofPlanes and Image-Edges hypotheses for 3D roof breakline elements are found. To accomplish this, the original 2D image edges are projected into the ground system again, but in difference to the first iteration, they are projected into the corresponding infinite roof plane (see Fig. 7a). Furthermore, the number of intersected edges can be reduced due to the information about the coarse position of the buildings. Then, an edge clustering process is applied for each roof plane separately, in order to locate the 3D roof plane boundaries precisely. The clustering is not limited to special roof types (any type of polygon can be reconstructed). The whole clustering process is based on two histograms – a histogram for the orientations and a histogram for the positions of all projected edges. Each projected edge contributes to the histograms depending on its length and its match with the PSE-Edges. The first step in the edge clustering is the generation of an edge orientation histogram and the extraction of local maxima in this histogram. These maxima describe the main directions of all projected edges. In the following each edge is assigned to one main direction and the position of the edges having the same main direction (plus/minus a few degrees) are examined. A local coordinate system is created and the intersections of the edges with the y-axis are computed and stored in the position histogram. For every local maximum in the position histogram which is above a certain threshold a line is drawn with the corresponding y-axis intersection and orientation. The result are a few infinite lines whose intersection represents the boundary of one roof plane (see Fig. 7b).



(a) (b) (c)

Figure 7: a) All "Image-Edges" intersected with both roof planes. b) Surrounding lines of roof planes after Clustering. c) Closed polygons of the roof plane boundaries.

Finally to clean-up roof topology all roof planes belonging to one building are fused by intersecting adjacent roof planes and snapping neighboring 3D polygon points. The result is a 3D representation of the roof structure and topology, consisting of a closed polygon surrounding the building and breaklines defining the transitions between single roof planes.

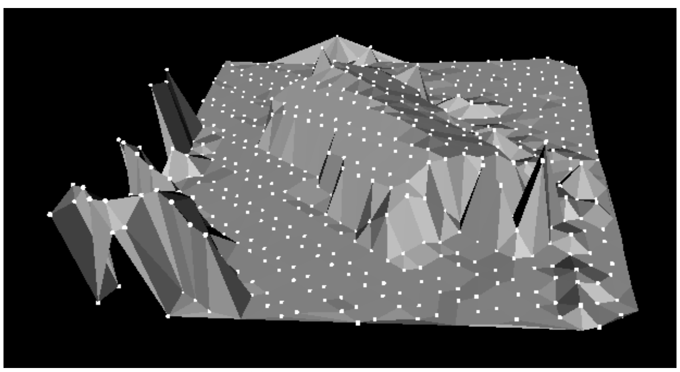
In a last step the LIDAR-TreeBlobs and Image-Blobs as well as the updated surrounding building polygon are mapped into a new Surface Class Image. After this information aggregation both the 3D outlines of each roof plane and the Surface Class Image serve as input information for the final surface estimation.

Final Surface Estimation: The goal of the final surface estimation is to generate a DTM without vegetation, but with buildings showing accurately modeled surfaces. Like in the preliminary surface estimation three different input components are used. First of all the raw LIDAR data, which is the genuine information and is the basis for the whole estimation process. It gets the same weight as in previous surface estimation process. As second input component the roof plane outline polygons are used in the estimation process. They represent the roof structure and topology in 3D and therefore get a high weight in the estimation process. Moreover, a slightly dilated duplication of the roof boundary is included to enforce the estimation of near-vertical walls. The final input component is the Surface Class Image. It is responsible for which parameter set is applied to which region. The result of the final surface estimation is illustrated in Fig. 8.

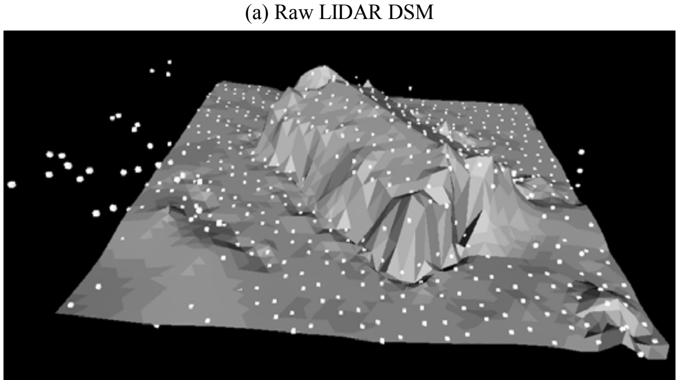
IV. Results and Discussion

In this paper a building reconstruction approach based on the fusion of LIDAR data and aerial imagery was presented. For the data fusion features had to be extracted from each data set separately. After the fusion of these features relevant building information could be provided for the surface estimation process. This estimation process was done iteratively, in a way that first the coarse building position was detected and then its exact shape was reconstructed. The approach was tested with different building types (flat roofs, gable roofs and combined roof structures) and provided good results also for relatively variable roof structures (see Fig. 9). However, further tests with buildings having a more complex roof structure should be made in order to confirm the generic reconstruction process.

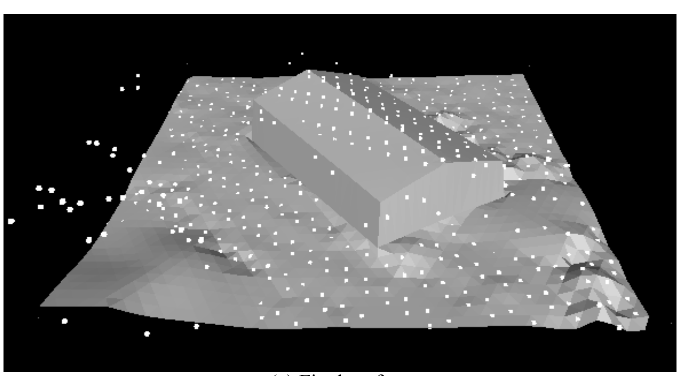
Future work to improve the presented approach could start at different points. A main task for example is the further development of the "topological clean-up" (cf. Sect. 3 B), because the reconstruction of a correct roof topology based on the found roof planes is crucial for the final surface estimation. At the moment at least a few edges of the roof structure have to be found in the aerial image in order to be able to reconstruct the building. Therefore further work is necessary to extract roof surface breaklines by making better use of LIDAR data. The drawback that edges have to be found in the imagery could also be overcome by applying the "Image-Edge" extraction to two or more aerial images. Up to now, only information extracted from LIDAR data and aerial imagery is fused in this approach. Additional information could also be extracted from other input data sets like e.g. GIS. It would be interesting to examine the value of an extended number of input data sets for the building reconstruction process.



(a) Raw LIDAR DSM



(b) Preliminary estimated surface.



(c) Final surface

Figure 8: From raw LIDAR DSM to the final building surfaces

ACKNOWLEDGEMENT

The approach presented in this paper was developed within the scope of the diploma thesis of Martin Huber. The work was supervised by Sanborn LLC, Colorado Springs, USA and the Chair of Photogrammetry and Remote Sensing at Technische Universität München. The support of both institutions is greatfully acknowledged.

REFERENCES

[1] Förstner, W. 1999: *3D-City Models: Automatic and Semiautomatic Acquisition Methods*. D. Fritsch and R. Spiller (eds.), Photogrammetric Week `99, Wichmann Verlag, pp. 291–303.

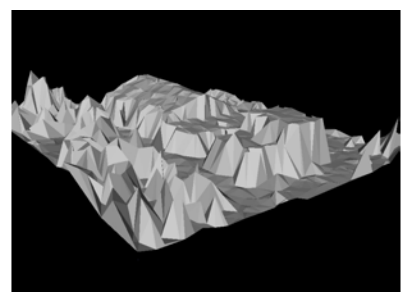
[2] Fuchs, C., Gülch, E. & Förstner, W. 1998: *OEEPE Survey on 3D-City Models*. OEEPE Official Publication No. 35.

[3] Brenner, C. 2000: *Towards Fully Automatic Generation of City Models*. International Archives of Photogrammetry & Remote Sensing, vol. 33, part B3, pp. 85-92, Amsterdam.

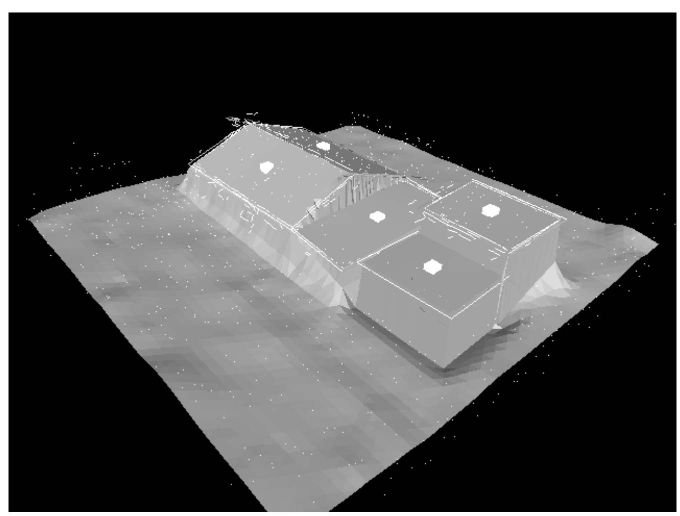
[4] Schickler, W. & Thorpe, A. 2001: *Surface estimation based on LIDAR*. Proceedings of the ASPRS Annual Conference St. Louis, Missouri, April 2001.

[5] Kraus, K. & Pfeifer, N. 1998: *Determination of terrain models in wooded areas with airborne laser scanner data*. ISPRS Journal of Photogrammetry & Remote Sensing 53, pp. 193-203.

(a)



(b)



(c)

Figure 9: Building example: a) Aerial image. b) Raw LIDAR point triangulation. c) Final estimated surface.