Conference Paper

Detection and reconstruction of buildings for a 3-D landscape model of Switzerland

Author(s):
Niederöst, Markus

Publication Date:
2000

Permanent Link:

https://doi.org/10.3929/ethz-a-004332774 →

Detection and reconstruction of buildings for a 3-D landscape model of Switzerland

Markus Niederöst
Institute of Geodesy and Photogrammetry (IGP), Swiss Federal Institute of Technology Zurich (ETHZ)
ETH Hönggerberg
CH-8093 Zurich, SWITZERLAND
e-mail: markus@geod.baug.ethz.ch
web: http://www.photogrammetry.ethz.ch/

KEYWORDS: Buildings, Blob detection, Unsupervised classification, Edge detection, 3-D reconstruction

ABSTRACT
The paper presents approaches both for the detection as well as the coarse reconstruction of buildings from middle scale aerial images (scale approx. 1:15'800) with the use of only one single stereo pair and no additional data from laser or infrared imagery. In a first step the buildings have to be approximatively located. In order to exclude the disturbing vegetation, an unsupervised classification is applied to separate vegetation areas from man-made image parts. Height values caused by vegetation are removed from a normalized surface model (nDSM) in which the buildings are detected by search of connected groups of rasterpoints above ground. Edge detection and determination of correspondences allow to calculate 3-D edges. With unsupervised K-means klassification the main building surfaces are determined and their spatial equation is calculated in a least squares adjustment using the reconstructed 3-D edges as well as elevation values from the digital surface model (DSM). Determination of four keypoints for each key surface and addition of walls leads to a coarse 3-D building model.

1. INTRODUCTION

Reconstruction of buildings is currently one of the main topics of the photogrammetric community, as such data is often required by public and private organizations in various applications (especially GIS-based ones) and for visualization. The work presented in this paper is part of the project ATOMI (Automated reconstruction of Topographic Objects from aerial images using vectorized Map Information), a cooperation between the Swiss Federal Institute of Topography, Bern (L+T) and the Institute of Geodesy and Photogrammetry (IGP) at ETH Zürich. Aim of project ATOMI is to update and correct approximate 2-D vector data (called VECTOR25) containing roads and buildings and to derive the height of the road centerlines and at least one representative height for each building. Roof outline details smaller than 2 m are not modelled. The motivation is to provide actual and precise building and road data for the production of a 3-dimensional landscape model of the whole area of Switzerland which is not a generalized cartographic one but corresponding to the real look of the landscape. For a more detailed overview on the project see (Eidenbenz et al, 2000). The investigations on roads are described in (Zhang/Baltsavias, 2000). Other approaches for building detection (e.g. with exclusive use of a DSM) and a reliable rough building reconstruction (based on orthophoto and nDSM data) with automatic evaluation of the results can be found in (Niederöst, 2000).

Approaches to solve the problem of building detection and reconstruction as well as the used kind of data are manifold. (Henricsson, 1996) used fourfold overlapping colour images (pixelsize in object space 7.5 cm). A procedure with even sixfold overlap is described in (Baillard et al, 1999) where the ground resolution is 8.5 cm. Both approaches take profit of the high ground resolution and the number of images and thus reduce the problem of edges that are not contained in some of the images. (Weidner, 1997) or (Maas, 1999) use very precise digital surface models from airborne laser scanning systems, while the classification procedures of (Haala/Walter, 1999) base on color infrared images which simplify the separation of man-made objects from vegetation. An overview of various approaches can be found in (Grün et al, 1995) and (Grün et al, 1997).

This paper shows the actual work done on the approximate detection of the buildings and - slightly apart from the aims of project ATOMI - a possible approach for the coarse 3-D reconstruction of buildings as well as it's limitations. Although approximate vector data for a rough location of most of the buildings would be available, it is not used at the present stage of investigations. As the vector data set never can be complete and always some buildings are missing or old buildings are contained that don't exist anymore, the aim is to develop a more general approach without a priori knowledge.

The main differences of the data used at present in project ATOMI to the work described by other authors is the use of a single stereo pair of medium scale images (scale 1:15'800) and the use of a DSM produced with photogrammetrical methods. No image information other than RGB is used.

2. BUILDING DETECTION AND RECONSTRUCTION (OVERVIEW)

The problems to be solved in the case of buildings can be separated in two parts. In a first step the rough location of the buildings is determined. If contained in the 2-D vector data set, the approximate building location could directly be used. In order to allow the addition of new buildings and the removal of those which do not exist anymore, blob detection in a digital surface model (DSM) is done. Combination with unsupervised classification allows to eliminate height values caused by vegetation. In the second step each detected building is reconstructed with edge detection followed by calculation of 3-D edges using corresponding edges in the two images. Color attributes and classification methods are used to select the edges and DSM points belonging to one building surface. The 3-D building surfaces can then be calculated in a least squares adjustment. Determination of keypoints for each main surface and the addition of walls allow to establish the 3-D model. Building outline and representative height(s) could now be derived.

3. INITIAL DATA

The used region is the village Hedingen south of Zurich, Switzerland. For the research phase the test area was reduced to a square of approx. 220 x 220 m^2. Main criteria for the selection of the test area were that it includes various types of buildings and some vegetation elements like trees (Fig. 1).

3.1 Image data

The color images of one stereo model (mean flight height above ground 4'800 m, focal length ~ 300 mm) were scanned with a resolution of 14 microns. At an image scale of 1:15'800 this resulted in a ground resolution of 0.22 m. The exterior orientation of the images was taken from an aerotriangulation.

Fig. 1: Test area as contained in the left color image

3.2 Digital terrain model

The digital terrain model (DTM) was provided by the Swiss Federal Institute of Topography (L+T). This so-called DHM25 was derived from the contour lines of topographic maps (scale 1:25'000) by interpolation. It is available for the whole area of Switzerland with a rasterwidth of 25 m. The accuracy of the DHM25 is around 1.5 m for the Swiss Plateau and approximatively 5 to 8 m for the alpine area.

3.3 Digital surface model (DSM)

Commercial software (Phodis by Zeiss) was used to generate a DSM (rasterwidth 1 m) of a part of the stereo model. Comparison of the achieved DSM result both from 14 micron and 28 micron images showed - with regard to the approximate building detection - no significant difference in quality.

4. APPROXIMATE BUILDING DETECTION

The blob detection for determination of approximate building locations was done by separation of regions above ground in the normalized DSM (nDSM = DSM minus DHM25) using a threshold. This allows to distinguish elements above ground from ground parts. Some blobs are just partially or not at all caused by man-made objects. Trees for example are not possible to separate from the buildings by exclusive use of height information. Hence unsupervised classification is used previous to the blob detection to eliminate height values caused by vegetation.

4.1 Determination of vegetation regions

The method for the separation of man-made objects from vegetation was unsupervised K-means classification with two input channels and two classes to be separated.

The left color image was used to extract the three separate color channels R (red), G (green) and B (blue). From those the following two input channels for the classification were calculated:

Channel 1: a* (redness-greenness) from the CIELAB color space which can be derived from R, G and B
Channel 2: arithmetic combination (G-R) / (G+R), inverted and scaled to the interval [0, 255]

Tab. 1: Input channels for determination of vegetation

The CIELAB cube root color coordinate system of the Commission Internationale de l'Eclairage (C.I.E.) was developed to provide a computationally simple measure of color in agreement with the Munsell color system. A unit color distance within the color solid should represent a perceptual color difference regardless of the particular pair of colors considered. Basically, L* is correlated with brightness, a* with redness-greenness and b* with yellowness-

blueness (Pratt, 1991). As vegetation usually is represented by green colors, it is obvious that the channel a* (redness-greenness) is chosen for the classification.

The arithmetic combination (G-R) / (G+R) showed the quality that vegetation is rather separated from man-made objects (Sibiryakov, 1996).

The classification result shows a sufficient separation of man-made objects from vegetation (Fig. 2).

Fig. 2: Classification result with 2 classes (black = man-made, white = vegetation)

4.2 Elimination of vegetation from DSM / nDSM

The result from the classification was used to mask the nDSM. Projection of the DSM height values back to the image of the classification result and check of a 11x11 pixel region allowed to calculate if a height value is caused by vegetation or not. All height values with more vegetation pixels than man-made pixels (in the 11x11 area) were assumed to be caused by vegetation and thus in the blob detection were handled as if they were ground points. The benefit of the elimination of vegetation is shown in Fig. 3 for the top right part of the test area.

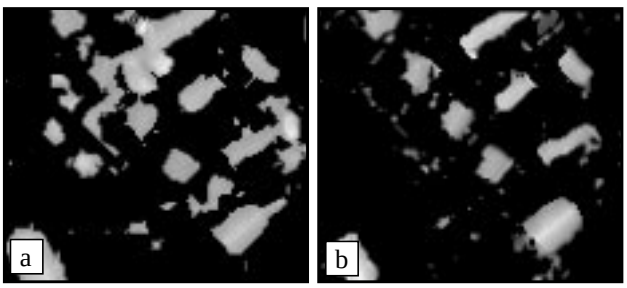


Fig. 3: Normalized DSM a) before elimination of vegetation and b) afterwards

4.3 Blob detection

The regions of interest (ROI) for building reconstruction were detected by search of adjacent height values higher than 2 m above ground, enclosed by lower areas. A minimum size threshold served to filter out small elements. To each region of interest a border of 5 m width was added. The result was a list of ROIs with the X, Y and Z coordinates of the top left corner and the width and the height of the ROI (all values in object space).

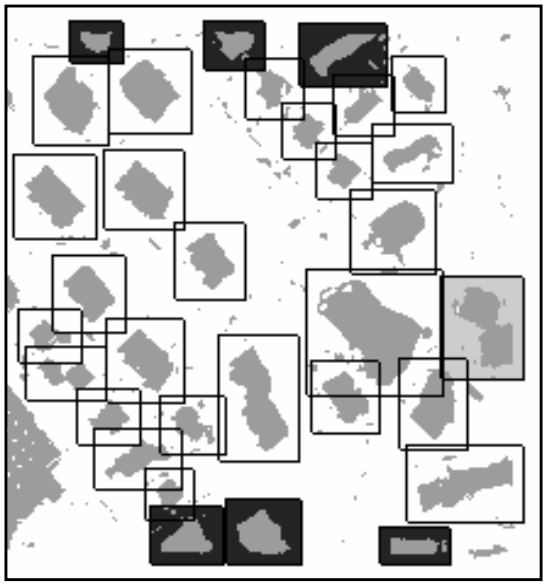


Fig. 4: Regions of interest after blob detection:
- rectangles represent ROIs
- dark grey = partially outside of test area,
- light grey = two close buildings, one partially outside of test area

Buildings that lie partially outside of the test area are not used in further steps. The case of the two buildings standing close together is processed (see Fig. 4).

In addition to the rectangular ROI the correct shape of the blob (projected to left image space) and the height values were saved in raster format for later use. This allows that only the centre building causing the blob is processed in later steps. Parts of other buildings or small high objects caused by errors that are located in the rectangular area can be excluded.

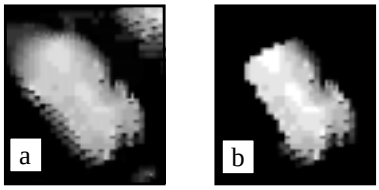


Fig. 5: DSM as greyvalues a) in the rectangular ROI, b) only for the exact blob

5. COARSE BUILDING RECONSTRUCTION

The proposed approach for coarse building reconstruction bases on the idea of the combination of seperately reconstructed roof parts. After the determination of the roof surfaces by classification methods their 3-D equation is determined using edge detection followed by the surface calculation with additional use of DSM points. Four keypoints are determined for each of those planes and the surfaces are combined. After the addition of walls the building model can be established.

5.1 Determination of characteristic roof surfaces (2-D)

To separate the roof surfaces, an unsupervised K-means classification with two channels is used. The chosen input channels are:

Channel 1: R (red) from RGB color space
Channel 2: L* (brightness) from CIELAB color space

Tab. 2: Input channels for separation of roof surfaces

The choice of channels R and L* was done after empirical evaluation. They needn't to allow the separation of particular classes (e.g. vegetation and man-made objects in 4.1) because the aim of this step is rather the separation of regions than the classification in order to know the type of the regions.

Content of *both* images is taken in account for the classification. The values for a particular channel are - for each pixel in left image space - calculated as average of the value in the left image and the value at the corresponding position in the right image, found by projection of the position in the left image onto the DSM and into the right image plane (Fig. 6). The influence of inaccuracies in the DSM can be disregarded under the assumption that - for radiometrically regular surfaces - adjacent pixels have similar values.

K-means classification is done with the calculated input channels to separate 5 classes (Fig. 7a). The number of 5 classes was empirically found and turned out to be applicable. Nevertheless the 5 classes sometimes do not allow a sufficient and satisfactory separation which is mainly caused by low contrast between two adjacent roof surfaces. On the other hand radiometrical changes on one particular roof part can cause an undesired segmentation.

The classification result is used to determine groups of adjacent pixels of the same class to separate roof segments (Fig. 7b). After detection of the start pixel of a region, the search of adjacent pixels inside the correct blob shape (see Fig. 5) is continued until no adjacent pixel of the same class is found.

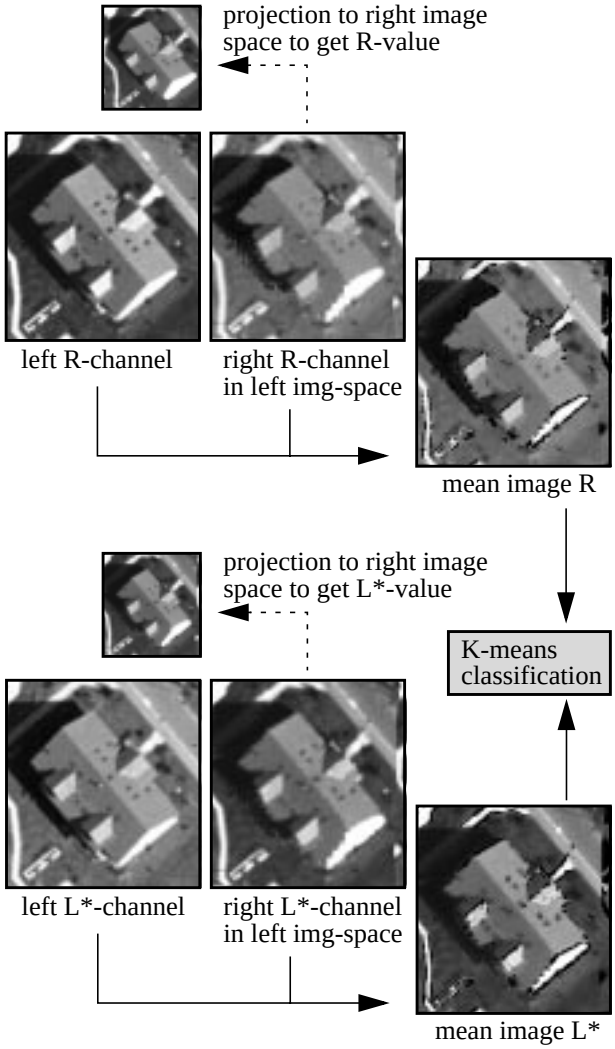


Fig. 6: Schematic view of the classification procedure

As the 3-D reconstruction of all roof surfaces from one single stereo model with image scale 1:15'800 and using just an optical DSM is not very probable, the regions are reduced to a sample of main roof surfaces which shall allow to represent the coarse building shape. Smaller surfaces are attached to the bigger ones by check of adjacencies. The postprocessing is done with morphological filtering using the operation *closing* (= dilatations followed by erosions) to fill gaps (Fig. 7c).

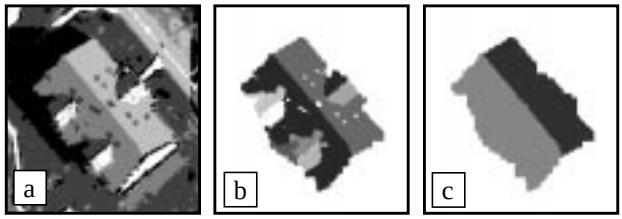


Fig. 7: Determination of main roof surfaces: a) result from K-means classification, b) determined roof surfaces, c) determined main roof surfaces

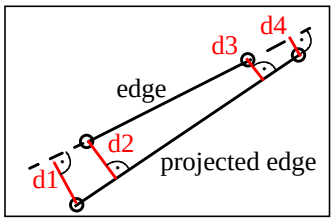
5.2 Roof edges

3-D edges were calculated with the assumption that a correct group of 3-D edges gives an idea of the building shape and allows the reconstruction.

Low level image processing was applied using the Canny operator followed by calculation of contour graphs (Henricsson, 1996). Only straight edges with a lengh > 7 pixels (1.5 m in object space) were accepted.

The search of corresponding edge pairs was done by projection of the 2-D edges onto the DSM and to the other image. The best candidate for each edge was found with the following conditions:

- the perpendicular distance of a startpoint or endpoint to the candidate edge projected from the other image (d2, d3) and vice versa (d1, d4) has to be small.



- the radiometric values of the flanking regions of both edges have to be similar; used channels are *L (brightness), *a (redness-greenness) and *b (yellowness-blueness) of the CIELAB color space).

Calculation of a 3-D surface equation using one edge (start and endpoint and the projection centre) and projection of its corresponding edge in the other image onto this surface leads to a 3-D edge (Fig. 8).

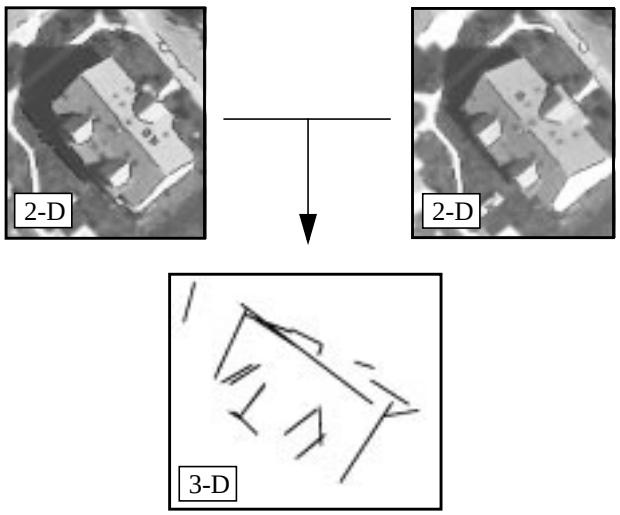


Fig. 8: Edges detected in the left and right image and reconstructed 3-D edges

The reconstruction of the building by the exclusive use of 3-D edges was in a few cases successful, when also for the human eye and brain it was possible to recognize the characteristic roof shape by viewing the reconstructed edges. But for many buildings the number of 3-D edges turned out to be insufficient.

5.3 Calculation of 3-D roof surfaces

The 3-D equations of the main roof surfaces are determined in a least squares adjustment. A surface can be written as

$$Z = a \cdot X + b \cdot Y + c \qquad \text{(Eqn. 1)}$$

Each start- and endpoint of a 3-D edge which is - in image space - adjacent to the roof surface forms one observation equation. Test results showed that not for each surface a sufficient number of reconstructed 3-D edges can be guaranteed. Thus a raster of DSM points (rasterwidth 0.66 m) which - projected to image space - lie on the actual surface is added (Fig. 9).

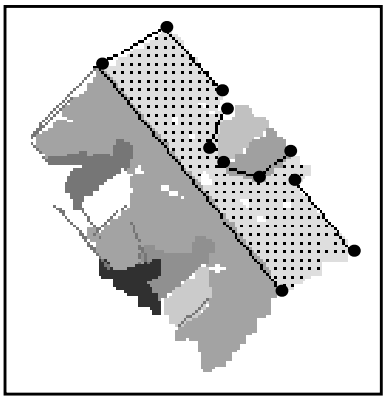


Fig. 9: Points for surface calculation projected to images space (endpoints of edges and DSM points)

The functional model for the least squares adjustment can now be formulated.

$$\begin{array}{lcl}
Z_{e1} & = & a\,X_{e1} + b\,Y_{e1} + C \\
\vdots & & \vdots \\
Z_{ei} & = & a\,X_{ei} + b\,Y_{ei} + C \\
\vdots & & \vdots \\
Z_{en} & = & a\,X_{en} + b\,Y_{en} + C \\
Z_{d1} & = & a\,X_{d1} + b\,Y_{d1} + C \\
\vdots & & \vdots \\
Z_{dj} & = & a\,X_{dj} + b\,Y_{dj} + C \\
\vdots & & \vdots \\
Z_{dm} & = & a\,X_{dm} + b\,Y_{dm} + C
\end{array} \qquad \text{(Eqn. 2)}$$

e = startpoint or endpoint of an edge, i = 1..n (number of edgepoints)
d = point on DSM, j = 1..m (number of DSM points)

Although the equation system is linear the constant number of 5 iterations is run through. After each iteration the weights of the observations are adapted according to their residuals (old weight divided by residual = new weight). This is done in order to eliminate the influence of blunders.

5.4 Reconstruction of the building

The regions derived from K-means classification do not allow a clear establishment of all points relevant for the correct roof surface. The reason is that misclassifications, occlusions and radiometric variability don't allow to determine the exact shape (all characteristic points) of each main roof surface.

Thus for the roof planes is assumed that they can be represented by four points. This is not generally valid, but has been chosen as procedure at this stage of research.

The 4 keypoints (2-D) are determined as follows:

1) The longest straight side of the region is detected by evaluation of all pixels along the outline of the surface and analysis of angles between the line formed by pixels n and n-3 to the line formed by pixels n and n+3, searching for the longest chain of angles which all are close to 180 degrees (threshold 20 degrees).

2) At both the start- and the endpoint of the first edge an additional edge is attached and rotated, searching for the position when it lies tangential to the region

3) The fourth edge is determined by parallel translation of the first edge until it lies tangential to the surface

4) The 2-D keypoints are located by intersection of the line equations of the four sides as described above.

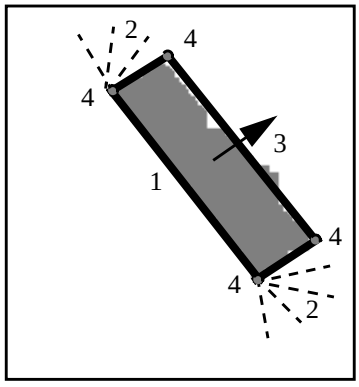


Fig. 10: Determination of keypoints

The 3-D coordinates of the 4 keypoints are then determined by projection of the 2-D keypoints onto the 3-D surface described in 5.3. Keypoints lying close to surface intersections are moved perpendicular onto the intersection line to close gaps and for model refinement.

With the availability of four 3-D keypoints for each separate main roof surface it's now possible to determine the coarse building model. Walls of the building are calculated by projection of the surface points onto the DHM25.

The result shows a coarse building model (VRML format) with the correct roof type (Fig. 11).

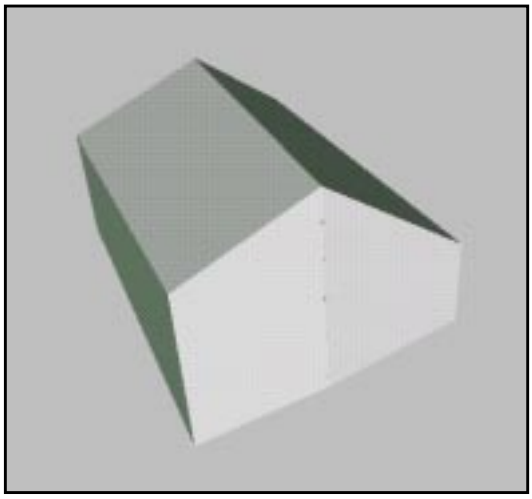


Fig. 11: Reconstructed building Nr. 25
(see also Fig. 14)

In the case that the classification procedure allows to separate the correct main roof surfaces, even the reconstruction of more complex buildings is possible (Fig. 12).

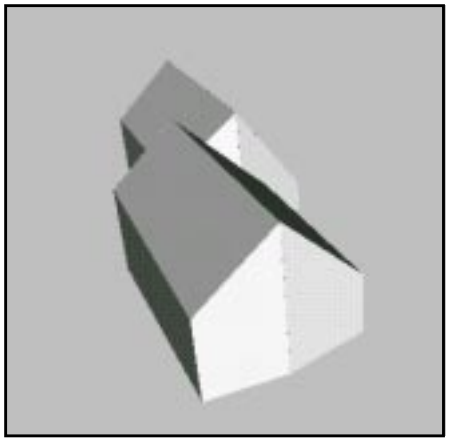


Fig. 12: Reconstructed building Nr. 147
(see also Fig. 14)

6. EVALUATION OF THE RESULTS

For the judgement of the procedures described above it is inevitable to do an objective analysis of the results.

The building detection was 100% successful. The approximate location of all 26 buildings fully situated inside the test area was determined. Considering the main task of project ATOMI which is the determination of the building outlines it should be possible - using those blobs combined with image content - to correct the initial vector data set which equals the update of a digital map.

Determination of 3-D edges was not a main research topic. Nevertheless, the detection of 2-D edges as described in (Henricsson, 1996) produces best possible results. The

search of corresponding edges using the DSM turned out to be applicable, but an extend for the use of epipolar lines (Zhang, 2000) could improve the reliability of the reconstructed edges. A major issue would be to find a solution for edges parallel to the epipolar line.

It is obvious that building reconstruction with only twofold overlap and a ground resolution of 0.22 m will not work in 100% of the cases, and especially the reliability of the found solutions is hard to be proved. That's why all reconstructed buildings in the test area are checked, and in the case of an insufficient result the reasons of failure are analyzed.

The developed procedures achieved nice results for the buildings 25, 28, 31, 55, 73, 78, 104, 121, 139, 147 and one part of the connected buildings 129 (see Fig. 14 and more detailed Fig. 11 and Fig. 12). They could directly be used to derive the building outline and representative roof heights for ridge line(s) and the roof outline.

The major number of failures are mainly caused by an insufficient result of the classification to separate the roof surfaces. In some cases the radiometric difference between two main roof surfaces was too small and even for the human eye hardly visible. Both surfaces were classified to belong to the same cluster (Fig. 13). For this reason just a flat roof was reconstructed (see building 76 in Fig. 14).

In other cases the radiometrical variety on one particular roof surface was to wide and thus the classification led to a segmentation of the surface into several parts (see building 93 in Fig. 14).

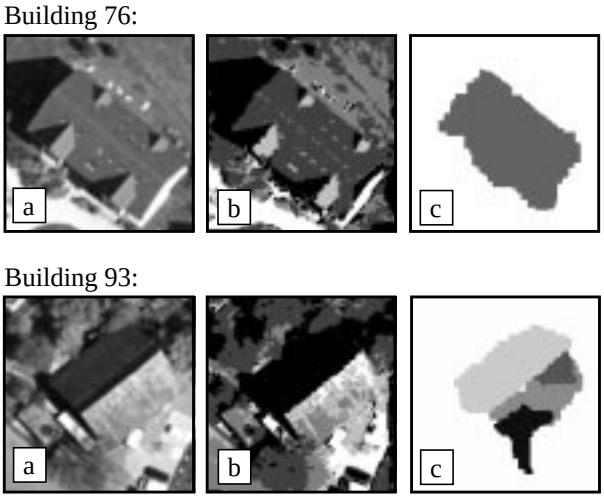


Fig. 13: Classification failure caused by radiometric properties: a) L*-channel, b) classification result c) main roof surfaces

A possible solution for the problem of the radiometric variability would be the use of a more precise DSM (e.g. laser

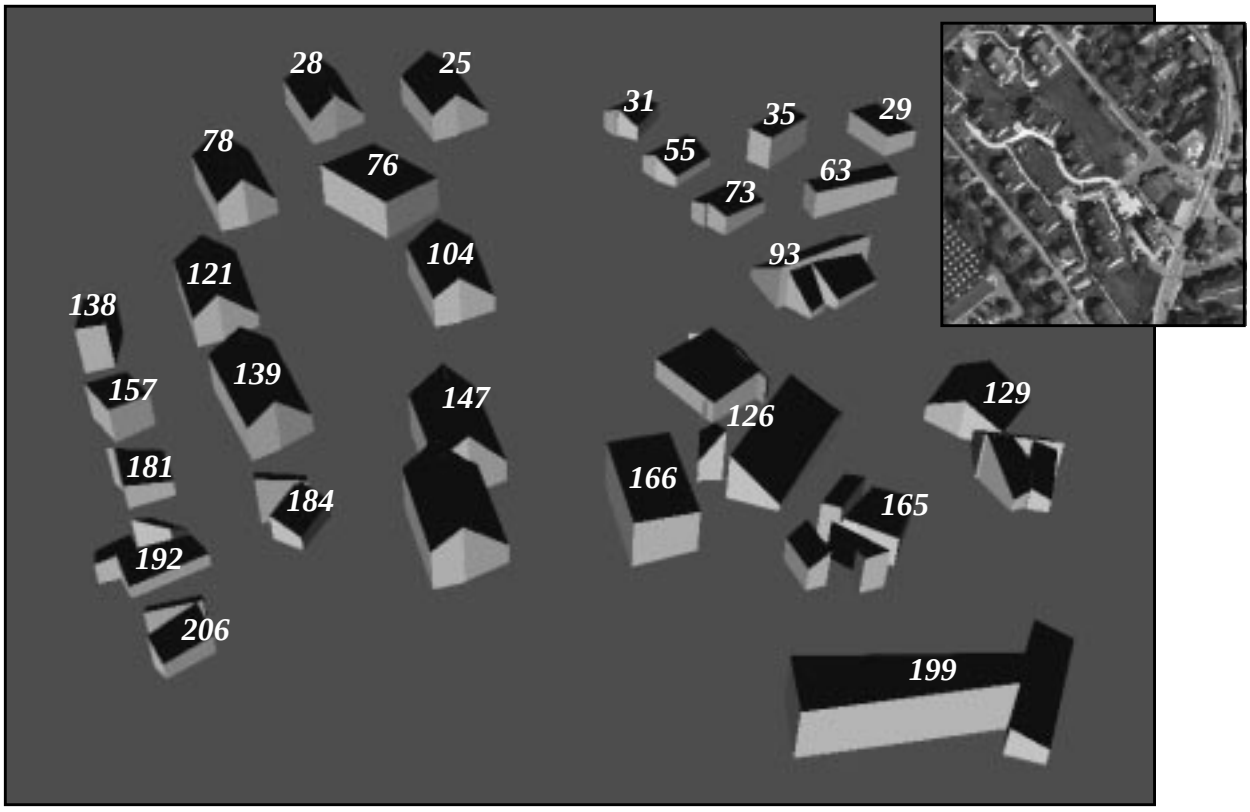


Fig. 14: Result of building detection and reconstruction for the whole test area

scanning data) which would allow the determination of the direction and slope of the height value gradients. The roof surface determination could then by supported by the search of surfaces with similar gradient direction and slope. In addition the laser DSM would allow to raise the accuracy of the reconstructed building.

DSMs from laser data will be used in project ATOMI since summer 2000. Tests will show how big the influence on the result really is.

For a final evaluation of the developped procedures, the results should be compared to a reference data set which was measured manually. This step is planned as soon as results with the use of laser data are generated.

7. CONCLUSIONS

In this paper methods for automated building location as well as coarse building reconstruction were presented. Though the use of only a single middle scale stereo model can cause problems if a specific feature is not visible in one of the images, it could be shown that a coarse building reconstruction from just one single pair of aerial images is possible. The result of the building detection with other image content will allow to fulfill the main tasks of project ATOMI, the correction of the approximate vector data and the reduction of the level of generalization. The determination of one or several characteristic building heights can then be done on the base of the DSM and/or the reconstructed buildings.

AKNOWLEDGEMENTS

The data for this research (digital aerial imagery and orientations, DHM25, VECTOR25) was provided by the Swiss Federal Institute of Topography, Bern.

REFERENCES

Baillard C., Schmid C., Zisserman A., Fitzgibbon A., 1999: Automatic line matching and 3D reconstruction of buildings from multiple views. IAPRS, Munich, Germany, Vol. 32, Part 3-2W5, pp. 69-80

Eidenbenz Ch., Käser Ch., Baltsavias E., 2000: ATOMI - Automated reconstruction of Topographic Objects from aerial images using vectorized Map Information. IAPRS, Amsterdam, Netherlands, Vol. XXXII

Gruen A., Kübler O., Agouris P. (Eds.), 1995: Automatic Extraction of Man-Made Objects from Aerial and Space Images. Birkhäuser Verlag, Basel

Gruen A., Baltsavias E., Henricsson O. (Eds.), 1997: Automatic Extraction of Man-Made Objects from Aerial and Space Images (II). Birkhäuser Verlag, Basel

Haala N., Walter V., 1999: Automatic classification of urban environments for database revision using lidar and color aerial imagery. IAPRS, Valladolid, Spain, Vol. 32, Part 7-4-3 W6, pp. 76-82

Henricsson O., 1996: Analysis of image structures using color attributes and similarity relations. Dissertation, Institute of Geodesy and Photogrammetry, ETH Zürich, Switzerland

Maas H.-G., 1999: Closed solutions for the determination of parametric building models from invariant moments of airborne laserscanner data. IAPRS, Munich, Germany, Vol. 32, Part 3-2W5, pp. 193-199

Niederöst M., 2000: Reliable reconstruction of buildings for digital map revision. IAPRS, Amsterdam, Netherlands, Vol. XXXII

Pratt W. K., 1991: Digital image processing. John Wiley & Sons Inc., New York

Sibiryakov A., 1996: House detection from aerial color images. Internal Report, Institute of Geodesy and Photogrammetry, ETH Zürich, Switzerland

Zhang Ch., Baltsavias E., 2000: Image analysis for 3-D edge extraction and knowledge-based road reconstruction. IAPRS, Amsterdam, Netherlands, Vol. XXXII