ROOF PLANE SEGMENTATION BY COMBINING MULTIPLE IMAGES AND POINT CLOUDS

Franz Rottensteiner

Institute of Photogrammetry and GeoInformation, Leibniz University Hannover,
Nienburger Strasse 1, 30167 Hannover, Germany – rottensteiner@ipi.uni-hannover.de

Commission III, WG III/4

KEY WORDS: Segmentation, Buildings, Data Fusion, Modelling, City Models

ABSTRACT:

A new method for roof plane detection using multiple aerial images and a point cloud is presented. It takes advantage of the fact that segmentation results for different views look different even if the same parameters are used for the original segmentation algorithm. The point cloud can be generated by image matching or by airborne laserscanning. Plane detection starts by a segmentation that is applied to each of the images. The point cloud is used to determine which image segments correspond to planes. The best plane according to a criterion is selected and matched with segments in the other images. Matching of segments requires a DSM generated from the point cloud, and it takes into account the occlusions in each image. This procedure is repeated until no more planes can be found. After that, planar segments are extracted based on region growing in the point cloud in areas of severe under-segmentation, and the multiple-image segmentation procedure is repeated. Finally, neighbouring regions found to be co-planar are merged. First results are presented for test site with up to nine-fold overlap. Our tests show that the method can deliver a good separation of roof planes under difficult circumstances, though the level of detail that can be achieved is limited by the resolution of the point cloud.

1. INTRODUCTION

1.1 Motivation and Goals

The 3D reconstruction of buildings has been an important topic of research in photogrammetry for almost two decades. Buildings can be reconstructed by a combination of parametric primitives that are fitted to the data or by polyhedral models, e.g. (Vosselman & Dijkman, 2001). If the second strategy is applied, it is necessary to detect roof planes in the input data which are then combined to reconstruct the buildings. Airborne laserscanner (ALS) data have frequently been used for that purpose. In ALS data having a resolution of 0.5 m-1.0 m, small roof planes cannot be found in the segmentation process, which restricts the level of detail (LoD) of the models. Furthermore, the delineation of roof planes at step edges was found to be difficult and not very precise in planimetry, e.g. (Vosselman & Dijkman, 2001). These problems can be overcome by using ALS data having a better resolution (Dorninger & Nothegger, 2007). The alternative is to extract the roof planes from digital aerial images based on a segmentation of these images. Aerial images are usually acquired at a higher spatial resolution than ALS data, so that more roof details are visible. Furthermore, the roof boundaries are represented better than in ALS data, so that the roof outlines should obtain a higher planimetric accuracy. However, image segmentation often does not perform very well in determining the roof planes in an aerial image. On the one hand, roof planes are often not separated correctly if the contrast between these planes is low, whereas on the other hand, roof planes are often split into several segments due to disturbances on the roof (Khoshelham, 2005; Song & Shan, 2008). Both phenomena, often referred to as under-segmentation and over-segmentation, respectively, usually occur at the same time in real-world scenes. Even if it was possible to tune the parameters of a segmentation algorithm so that it delivered a perfect delineation of the roof planes in a particular scene, there would be no guarantee that these parameter values could be transferred to other scenes. One way to make segmentation more robust with respect to parameter selection is to use additional information. In the past this was done by combining the results of the segmentation of a single digital image with a 3D point cloud from ALS (Khoshelham, 2005) or image matching (Drauschke et al., 2009). In the latter case, multiple images were available, but they were only used for matching. In this paper we present a new method for roof plane detection that combines the segmentation results of *multiple* aerial images with a 3D point cloud. Using this method a good separation of roof planes can be achieved even if there is not a single aerial image in which all planes are separated properly by the original segmentation algorithm.

1.2 Related Work

Dorniger and Nothegger (2007) have shown how excellent segmentation results can be achieved based on point clouds alone if the point density is very high. However, a resolution of 20 points / m² is not standard for ALS point clouds today and requires considerable costs for data acquisition. If applied to high-resolution point clouds from matching, their method resulted in a lower LoD due to noise in the point cloud. Song and Shan (2008) proposed a method for the segmentation of single colour images. They used an active contour model for delineating buildings and then generate an image called the J-image representing the local homogeneity of colour vectors inside the building regions. The J-image is used as the basis of a watershed segmentation to obtain the roof planes. This method can result in a very precise delineation of roof planes, but the separation of roof planes failed in case of poor contrast. Drauschke (2009) has shown how the results of watershed segmentation can be improved by integrating segmentation results at different scales based on an irregular pyramid. While maintaining the level of generalization that is the result of a higher degree of smoothing of the input image, the procedure results in a very precise delineation of the segment boundaries, obtained from the segmentation results achieved using a low degree of smoothing. However, selecting the regions corresponding to roof planes in the irregular pyramid has not yet been achieved (Drauschke, 2009).

All the algorithms discussed so far use either image or height data. Of course, these data sources can be combined. Ameri and Fritsch (2000) determined seed regions for region growing based on a local curvature analysis in a Digital Surface Model (DSM). These seed regions were used for region growing in a digital aerial image. The segmentation results were affected by the fact that roof planes having similar grey levels could not be separated correctly. In our previous work (Rottensteiner et al., 2004) we used region growing to detect roof planes in a DSM generated from ALS data. This was followed by a second region growing process taking into account multiple aerial images. Whereas the roof planes could be delineated more precisely thanks to using the images, the number of roof planes and thus the structure of the resulting building model were solely determined by the segmentation of the ALS data. Thus, the images did not add any new roof details to the model. Khoshelham (2005) also combined image and ALS data for roof plane segmentation. He carried out a watershed segmentation of a digital orthophoto and determined planes from the ALS data in each of the image segments using RANSAC. In order to deal with under-segmentation, image segments for which RANSAC delivered more than one plane were split into smaller parts, each representing a single roof plane. Neighbouring segments found to be co-planar were merged. RANSAC requires a relatively large number of points per roof plane to work, which limits the LoD that can be achieved by this method (Khoshelham, 2005). Drauschke et al. (2009) combined multiple aerial images for segmentation. Watershed segmentation was applied to one of the images, whereas the other images were merely used to obtain a point cloud from image matching. The point cloud was used to compute an adjusting plane for each segment, and neighbouring segments found to be co-planar were merged. Since no efforts are made to solve the under-segmentation problem, the method is not able to separate roof planes having a low contrast in the image used for watershed segmentation.

In this paper we want to combine the results of watershed segmentations of multiple aerial images and a point cloud for the improved segmentation of roof planes. This is motivated by the fact that roof planes having a low contrast in one image might well show good contrast in another view. An obvious example is a roof plane separated from its neighbour by a step edge. If the two planes have similar colour, they might not be separable in a nadir view, but then a wall will be visible in the stereo partner, so that the planes are easily separated. The point cloud is needed to check whether a segment corresponds to a plane and to match segments from different images. The new method is presented in detail in this paper along with first results that show its potential for roof plane segmentation.

1.3 Test Data

The German Association for Photogrammetry and Remote Sensing (DGPF) has acquired a test data set over the town of Vaihingen (Germany) in order to evaluate digital aerial camera systems. It consists of several blocks of vertical images acquired by various digital aerial camera systems at two resolutions. There is also an ALS point cloud with a density of 4-6 points / m^2. A detailed description of the data can be found in (Cramer & Haala, 2009). We used one of the DMC blocks and the ALS point cloud. The images are 16 bit pan-sharpened colour infrared images with a ground sampling distance (GSD) of 8 cm (flying height: 800 m, focal length: 120 mm). The georeferencing accuracy without compensating for systematic errors is about 1 pixel. The nominal forward and side laps of the images are 65% and 60%, respectively. As a consequence, each building in the centre of Vaihingen is visible in 6-9 images. The

ALS point cloud consists of several overlapping strips. Multiple pulses were recorded. The point cloud was pre-processed to compensate for systematic offsets between the strips (Haala et al., 2010).

2. ROOF PLANE SEGMENTATION

2.1 Overview

The input for our method consists of at least two aerial images with their orientation parameters and a point cloud. The point cloud can be generated by image matching or by ALS. Typically, it will have a lower resolution than the aerial images. We assume the accuracy of the points, represented by the standard deviations σ_{XY} of the planimetric coordinates and σ_Z of the height, to be known. Finally, approximate building boundaries and heights are required to define regions of interest.

In a pre-processing stage the point cloud is used to interpolate a DSM grid with a resolution corresponding to the GSD of the images. After that, an occlusion mask is derived for each image from its orientation parameters and the DSM. Each occlusion mask shows the areas in object space that are not visible in the corresponding image. Estimating the approximate floor height Z_{floor} as the 1% quantile of the point heights and using a pre-defined threshold ΔZ_{min} for the minimum building height above ground, an approximate digital terrain model (DTM) is interpolated from points with heights $Z < Z_{floor} + \Delta Z_{min}$.

Multi-image segmentation starts with a watershed segmentation of each of the aerial images. The resulting label images are projected into object space, taking into account the occlusion masks. For each of the segments in each of the projected label images, an adjusting plane is determined from all the off-terrain points inside the segment, using robust estimation to eliminate outliers. The best plane according to a criterion taking into account the goodness of fit and the number of points inside the plane is selected as an initial roof plane segment. This roof plane segment is matched with segments from the other images in order to expand it as much as possible by overlapping co-planar segments from other images. The resulting roof plane is represented by a segment in a *"combined label image"* defined in the (X,Y) plane of object space. The area covered by this segment is set to zero in the projected label images, and the procedure of computing planes, selecting the best segment and segment matching is repeated until no more planar segment with a minimum number NP_{min} of points can be found.

After the combined segmentation, neighbouring segments found to be co-planar by a statistical test (Rottensteiner et al., 2005) are merged. At this stage, some roof planes may remain undetected due to very low contrast. The point cloud is analysed to detect new planar segments by region growing. As a consequence, image segments corresponding to multiple roof planes are split, so that a second multi-image segmentation procedure can be used to detect the remaining roof planes. Finally, small gaps between the detected roof planes are filled.

2.2 Segmentation of the Individual Images

We apply watershed segmentation to the original images in a similar way as Drauschke et al. (2009). The basis for watershed segmentation is given by a an image $h(x,y)$ generated from the intensity image $I(x,y)$ according to

$$h(x,y) = \max(||\nabla I(x,y)||, k \cdot \sigma_I) - k \cdot \sigma_I \qquad (1)$$

where $||\nabla I(x,y)||$ is the norm of the gradient of $I(x,y)$, σ_I is the standard deviation of the intensity determined by the method

described in (Brügelmann & Förstner, 1992), and k is a parameter that controls the degree to which noise effects are eliminated (Drauschke et al., 2009). The input function for the watershed segmentation is a smoothed version of $h(x,y)$. Smoothing consists of N_s iterations of a convolution with a 3 x 3 binomial filter. Thus, the main control parameters of the segmentation are k and N_s.

The results of watershed segmentation are projected into object space. For that purpose, a grid having approximately the same GSD as the aerial images is defined in the (X,Y) plane in object space. The grid point heights are taken from the DSM, and each of the 3D points thus obtained is back-projected to the aerial image. The label found at the back-projected position in image space is written into the projected label image. After that, all pixels found to be occluded are set to zero. This procedure requires the occlusion masks generated from the DSM and the orientation parameters by ray tracing. After projecting the segmentation results to object space, an average height above the terrain is computed for each of the projected segments from the differences between the DSM and the DTM of all pixels inside the segment. Segments having an average height above the terrain smaller than the minimum building height ΔZ_{min} are discarded. Fig. 1 presents an example for the segmentation process. The upper row shows a building of medium complexity in three images. In the second row, the segmentation results are shown. There is a considerable over-segmentation of the roof parts lit by direct sunlight, whereas the shaded half of the building is under-segmented. It consists of two planes that intersect at an obtuse angle and that both enclose a large dormer. The two main roof planes in the shadow are not separated in any of altogether eight views available for this particular building, nor are the two dormers separated correctly from their enclosing planes. Fig. 1 also indicates the problems related to tuning the segmentation parameters. To avoid the under-segmentation in the shadow areas, a lower degree of smoothing or noise suppression would be required, but this would result in a gross over-segmentation of the sun-lit parts. On the other hand, using a higher degree of smoothing would reduce the degree of over-segmentation, but it would also result in even more planes being merged due to poor contrast.

By projecting the label images into object space, image segments corresponding to walls are nearly eliminated (cf. last row of Fig. 1). Additionally, some segments are split by the visibility analysis. For instance, in the original segmentation of the rightmost image (second row of Fig. 1), segment A covers both the shaded part of the roof and a shadow area on the ground. Projecting the segment to object space and setting the occluded areas to zero in the resulting label image separates the roof from the ground (segments A_1 and A_2 in the last row of Fig. 1). Fig. 1 also shows that the segmentation results achieved for several images are quite different. This is why, rather than tuning the segmentation algorithm so that the best possible segmentation is obtained in a "master" image, we combine the segmentation results achieved for multiple images, using a trade-off in the selection of the segmentation parameters.

2.3 Iterative Generation of Roof Plane Segments

The combination of the original segmentation results is an iterative process. In each iteration we try to find the most suitable planar segment among all the segments in all images. This planar segment is matched with segments from all the other images, and the combined segment is added to the list of planar segments. This process is repeated until no further planar segments segment can be found.
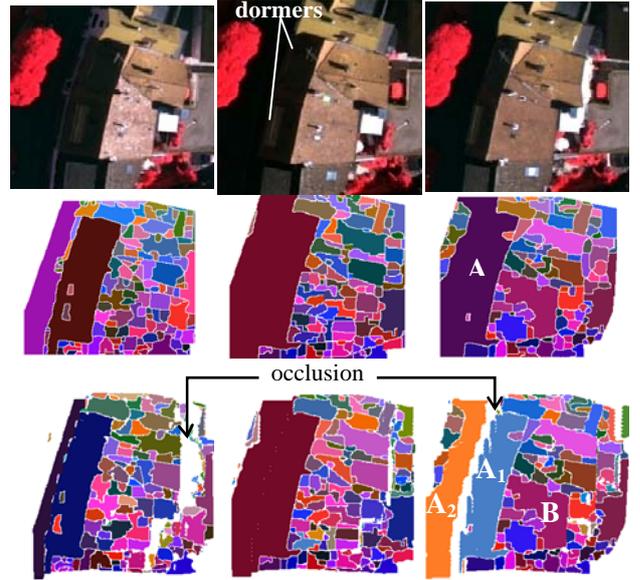

Figure 1. Upper row: original images; two dormers are indicated in one image. Second row: results of watershed segmentation using $k = 2$ and $N_s = 5$. Last row: segmentation results projected to object space. Two areas where occlusions occurred are marked.

2.3.1 Finding the Best Plane: Each iteration starts with the determination of adjusting planes from the off-terrain points for all the segments in all projected label images. Off-terrain points are points being at least ΔZ_{min} above the DTM. In each projected label image, each off-terrain point is assigned to the segment it is situated in. An adjusting plane is computed for each segment by a least squares adjustment minimizing the square sum of the distances between the points and the planes. Robust estimation by reweighting the observations based on the size of the residuals (Förstner & Wrobel, 2004) is carried out in order to remove outliers (e.g. points on chimneys). Robust estimation proceeds until the maximum distance is found not to differ significantly from zero by a statistical test (Rottensteiner et al., 2005). However, it is also terminated when more than a predefined maximum number of outliers Out_{max} are eliminated or when there remain less than a predefined number NP_{min} of points in the segment. Any segment is accepted as a valid plane if it fulfils three criteria: (1) The plane contains at least NP_{min} points, (2) there are no more outliers (i.e., the largest residual does not differ significantly from zero), and (3) the points inside the segments all lie on a plane. The third criterion can also be formulated as a statistical test. As we know the accuracy of the point cloud, the r.m.s. error s_0 of the weight unit of planar adjustment can be tested according to whether it can be "explained" by the uncertainty of the points. Representing the uncertainty of the points by the standard deviation σ_Z of the height, this corresponds to comparing s_0 to a threshold s_{max}:

$$s_{max} = \sqrt{\chi^2_{1-\alpha;red} / red} \cdot \sigma_Z \qquad (2)$$

In Equ. 2, $red = NP - 3$ is the redundancy, where NP is the number of points in the plane, and $\chi^2_{1-\alpha; red}$ is the $(1-\alpha)$ - quantile of the χ^2_{red} distribution. Using Equ. 2 rather than a fixed threshold for s_0 takes into account the fact that s_0 is more uncertain for small NP. All valid planes are ranked according to a score function $score(NP, s_0)$:

$$score(NP, s_0) = NP / \max(s_0, \sigma_Z/2) \qquad (3)$$

The score function favours planes that contain a large number NP of points while having a good planar fit as indicated by s_0.

Dividing NP by $\max(s_0, \sigma_Z/2)$ avoids a numerical overflow for very small values of s_0 and means that planes with $s_0 < \sigma_Z/2$ are effectively ranked by NP. The best plane is the plane achieving the highest value of the score function in Equ. 3. It is selected for further processing.

2.3.2 Matching of Segments: The best plane was selected in one of the projected label images. Due to uncertainties of the DSM at step edges, the boundaries of the projected segment may not be represented well. Thus, the original image segment corresponding to the plane is again projected into object space, this time using the adjusting plane to obtain the heights, so that the segment boundaries should be more precise. After that, we check whether all parts of the projected image segment receive sufficient support from the point cloud. Fig. 2 shows a situation that should be avoided: the segment only contains points in its left half, whereas there are no points in the right part. This may happen if a segment covers both a roof plane and a part of the ground, but has a good planar fit because only off-terrain points are used for determining the adjusting plane. Each point assigned to the plane is supposed to be representative for a circular area in the (X,Y) plane whose radius r is chosen to be slightly (10%) larger than the average point spacing of the point cloud. Any pixel assigned to the segment that is further away than r from its nearest point is erased in the label image. This might split the segment into two or more parts; in this case, the largest part is maintained. Due to cutting off at a distance r from the nearest point, the boundary of the segment is not very well defined at the locations where it had to be cut off (Fig. 2).
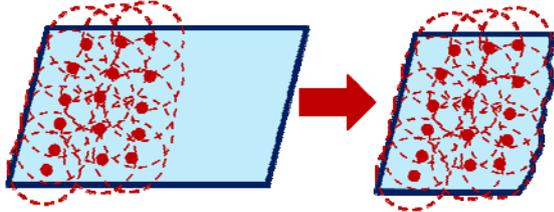

Figure 2. Left: A planar segment projected to object space. Red dots: points assigned to the plane. The circles are centred at the points and correspond to the area for which the point is representative. Right: the segment is cut off.

Having thus improved the shape of the planar segment, the actual matching process is carried out. Matching candidates are searched for in each image except the one from which the segment was taken. Each valid plane found in any of these projected label images (cf. Section 2.3.1) that overlaps with the planar segment is considered to be a matching candidate. If a matching candidate and the segment to be matched are found not to be co-planar based on a statistical test, the matching candidate defines an area that should not be merged with the plane. All such areas are marked in a binary "non-co-planarity-image". Otherwise, the candidate is accepted if there are a sufficient number N_{OL} of points that have been assigned to both planes; the latter criterion enforces that there are actually points that are co-incident with both planes. Once a candidate has been accepted, the original image label corresponding to the accepted candidate is re-projected to object space using the adjusting plane, and it is checked for support from the point cloud in the same way as the segment to be matched (Fig. 2). Matching is guided by an "overlap counter", i.e. an image that counts the number of overlapping segments at any position, including the segment to be matched and the accepted matching candidates. This overlap counter is initialised by the segment to be matched, and it is incremented at all positions covered by any matching candidate after re-projection to the adjusting plane.

The left part of Fig. 3 shows the overlap counter for the best segment (segment B) in Fig. 1. The total number of overlapping images in this case is eight. There is up to five-fold overlap, partly also with relatively small matching segments. In the centre of Fig. 3, the non-co-planarity-image is shown. From these two images, the final roof segment is generated. Firstly, all pixels covered by at least two segments are accepted (all except the red ones in Fig. 3). Pixels only covered by one segment (red pixels in Fig. 3) are only accepted if they do not belong to a non co-planar area (black pixels in the middle part of Fig. 3). Thus, multiple overlap can override the non-co-planarity criterion. Finally, binary morphological opening using a small square structural element is used to smooth the boundaries of the resulting segment (right part of Fig. 3).

Figure 3. Left: overlap count for the best planar segment (segment B) in Fig. 1. Red / yellow / green / blue / cyan: 1 / 2 / 3 / 4 / 5 overlapping segments. Centre: regions belonging to non-co-planar matching candidates (black). Right: final roof segment.

2.3.3 Accepting the Matched Segment and Iteration: The segment generated in the way described above is added to the combined label image of planar segments, and the segment's point list is updated to contain all points within the segment that are coincident with the adjusting plane. After that, the plane parameters are recomputed. Finally, the area covered by the new segment is set to zero in the projected label images (cf. Fig. 4, left). After that, the procedure of determining the best planar segment, matching of planes and adding the resulting segment is repeated until no more segments can be found. Erasing the accepted segment may change the structure of the projected label images. Most importantly, image segments that spanned several roof planes in some of the images (which would have resulted in a poor planar fit) may be separated in these images if the one of the roof planes could be detected based on another image at some stage in the iteration process.

2.3.4 Merging Co-planar Segments: The generation of roof segments is terminated when no more planar segments having at least NP_{min} points can be found in the data. The results of multi-image segmentation are represented by the combined label image of all roof segments. This label image is analysed for neighbourhood relations, and planes found to be co-planar based on a statistical test are merged. As the order of merging has a significant impact on the merging results, neighbouring planes are merged in the order of the combined r.m.s. error of the planar fit. The central part of Fig. 4 shows the combined label image for the building in Fig. 1. The rightmost label image in Fig. 4 shows the results after merging of co-planar segments.

2.3.5 Region Growing and Completion of the Segmentation: In case of large contrast variations in the images, smoothing will have the effect that some larger roof planes may not be separable by multi-image segmentation. This is why we check the point cloud for any planes that have been missed. For each off-terrain point not yet assigned to any roof plane we get its N_{nn} nearest neighbours in the point cloud. If neither of these points has been assigned to a roof plane, we determine the adjusting plane through these points and check whether it can be accepted as a plane in the way described in Section 2.3.1. If no plane is accepted, the process is terminated. Otherwise, the plane having

the best planar fit (indicated by s_0) is accepted as a seed region for region growing. Points not yet assigned to any other plane are added to the seed region if they are found to be coincident to the plane based on a statistical test. The resulting point set is analysed whether it corresponds to a connected segment in object space. If it is split into several segments, the segment containing the largest number of the original N_{nn} nearest neighbours is maintained. The parameters of the adjusting plane are determined from all points contained in the segment, and the planar segment is added to the combined label image. This procedure is repeated until no new plane can be found.
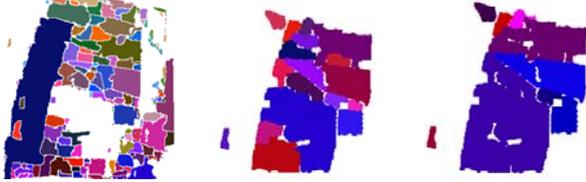

Figure 4. Left: the left projected label image from Fig. 1 without the segment from Fig. 3 and without ground segments. Centre: multi-image segmentation. Right: results after merging co-planar segments.

By erasing the segments generated by region growing in the projected label images, image segments corresponding to several roof planes may be separated. Thus it makes sense to repeat multi-image segmentation, this time using $NP_{min} = 3$ so that also very small segments can be detected. Finally, we try to close small gaps in the segmentation results by checking whether there are points that have not yet been assigned to any segment, but are coincident to one of the planar segments in their vicinity. These points are assigned to the nearest segment in terms of the point's distance from the plane. Again, co-planar segments are merged. At this stage, the segmentation results may contain segments that correspond to roofs of neighbouring buildings if they are close together, or they may even contain segments on vegetation or other high objects in the vicinity of a building. We eliminate planes having an overlap smaller than O_{min} of the segment's area with the region of interest defined by the approximation. No further classification is carried out.

The left part of Fig. 5 shows the results of the region growing process for the building in Fig. 1. Three new planar segments were added by the region growing process, two corresponding to the dormers and one that merges two roof planes where they intersect at an obtuse angle. After that, the second multiple image segmentation process can find a few more roof planes, because they have been cut off from the large image segments in the shadow areas (central part of Fig. 5). The rightmost part in Fig. 5 shows the final segmentation results after filling small gaps and merging co-planar segments.

3. RESULTS AND DISCUSSION

In order to test the method described in this paper, it was used to detect the roof planes of seven buildings in the historic centre of Vaihingen, using the data described in Section 1.3. The approximate building outlines were generated by the building detection algorithm described in (Rottensteiner et al., 2004). The planimetric accuracy of these outlines is about ±2 m. The buildings in this area show medium to high complexity. The number of images a building was visible in varied between six and nine, because the buildings are situated in the overlapping area of three strips. Tab. 1 shows the parameter settings used in our experiments. The results of our method, one of the images used for achieving those results, and a reference segmentation based on photogrammetric plotting are shown in Fig. 6.

Figure 5. Left: results after adding new segments based on the point cloud to the results in Fig. 4. Centre: results after the second segmentation. Right: results after merging co-planar segments and filling small gaps.
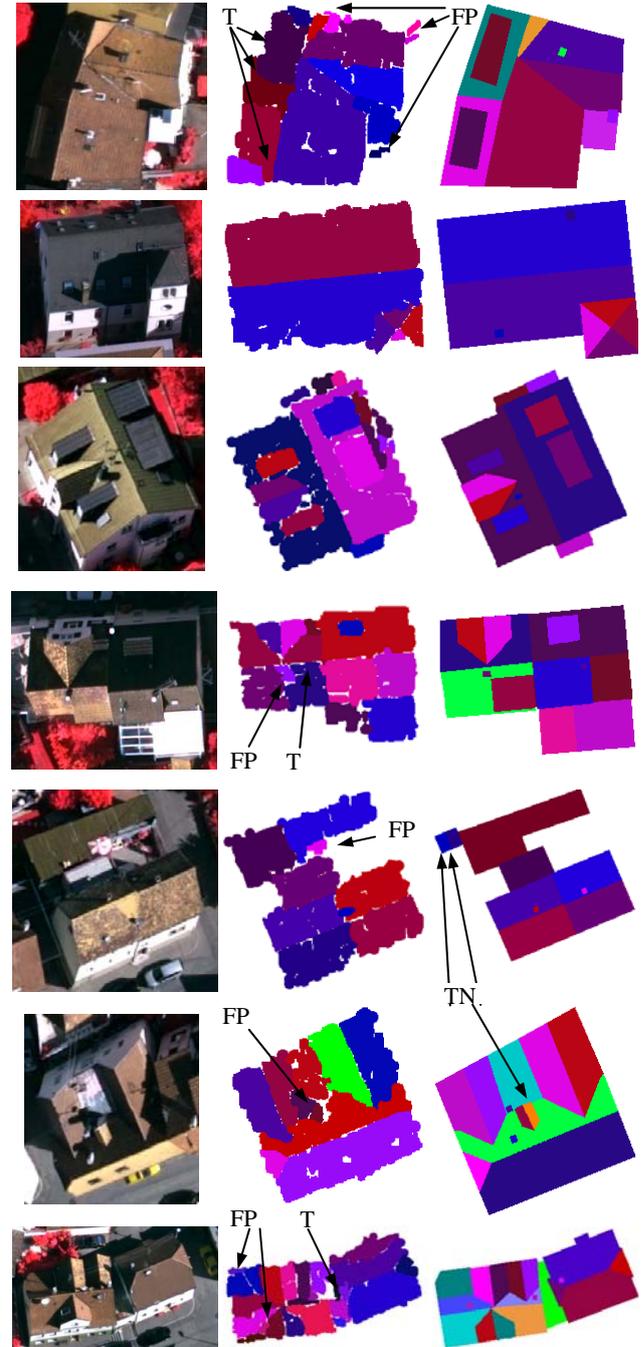

Figure 6. Left: one of the aerial images. Centre: segmentation results. Right: reference. FP: False positives. TN: True negatives (missed planes). T: Topological errors.

Fig. 6 shows that multi-image segmentation does a good job in separating the individual roof planes. This is confirmed by a

comparison of the results with the reference. The reference contains 85 segments. However, 15 of these segments represent chimneys and are smaller than 1 m^2. None of these chimneys could be found, but of the remaining 70 planes, 58 (82.9%) are detected correctly (based on a visual inspection). Another 10 planes (14.3%) are detected, but are split into two or more segments by our algorithm, and only three planes are missed. One of them is a small plane (1.5 m^2) that belongs to a dormer, whereas the other two correspond to a small building structure that was missed in the detection and thus is outside the region of interest for segmentation. Eight planes are affected by topological errors in the sense that a part of one plane is erroneously added to another plane so that the neighbourhood relations are affected (the splitting of planes can also be considered a topological error). Most of them affect dormers that intersect the main roof plane: the part of the roof plane that separates the dormer from the ridge is assigned to the dormer. Most of these errors occur with the building used for Figures 1 – 5, which is the only building where planes had to be added to the segmentation based on the point cloud (cf. Section 2.4). There are five false positive planes in the interior of the buildings. They are small structures at the transitions between several roof planes. There are four false positives outside the buildings. One of them is a parasol, the others are artefacts that correspond to vegetation and could possibly be removed based on their radiometric content. The delineation of the detected roof planes in Fig. 6 is not very precise yet. The deviations are in the order of magnitude of 3-5 image pixels and are clearly influenced by the resolution of the point cloud. In the future, the delineation of the roof planes shall be improved by integrating image edges in a way similar to (Rottensteiner et al., 2004).

σ_{XY}	σ_Z	ΔZ_{min}	k	N_s	α	NP_{min}	N_{nn}	N_{OL}	O_{min}
0.15 m	0.075 m	2.0 m	2.0	5	1%	4	16	5%	33.3%

Table 1. Parameters used in the experiments. The symbols are explained in the text. NP_{min} is given for the first segmentation process. N_{OL} is 5% of the number of points in either of the segments.

4. CONCLUSIONS

We have presented a new method for roof plane detection based on the segmentation of multiple aerial images and a point cloud. The method makes use of the observation that segmentation results differ over various images, so that a combination of several segmentations may result in a better separation of roof planes than a segmentation of a single image. The method was applied to detect the roofs of several complex buildings in a densely built-up historic town centre, based on images having a resolution of 8 cm and an ALS point cloud with an average point spacing of about 0.5 m. The results show that the method is capable of detecting most of the roof planes correctly with only a few false positives. The accuracy of the delineation corresponds to the average resolution of the point cloud and requires improvement for building reconstruction, e.g. by integrating image edges. The LoD of the segmentation is restricted by the resolution of the point cloud, because a minimum of three points is required for an image segment to be considered a candidate for a plane. This could be improved by integrating 3D edges derived from multiple-view matching, because only one further point (or an additional 3D edge) would be sufficient to support a plane hypothesis. It has to be noted that the current procedure essentially works in 2.5D because the projection of the segmentation results into the (X,Y) plane in object space is used for segment matching and because the combined segmentation is also represented in the (X,Y) plane.

This restriction could be overcome by defining and matching the planar segments directly in image space, in which case the images would be connected via the object planes. However, this would require a specific visibility analysis for any pair of images involved in the process. Finally, the method still needs to be verified for point clouds generated by image matching.

REFERENCES

Ameri, B., Fritsch, D., 2000. Automatic 3D building reconstruction using plane-roof structures. In: Proc. ASPRS Spring meeting, on CD-ROM.

Brügelmann, R., Förstner, W., 1992. Noise estimation for color edge extraction. In: W. Förstner, S. Ruwiedel (eds.), Robust Computer Vision, Wichmann, Karlsruhe, pp. 90-107.

Cramer, M., Haala, N., 2009. DGPF Project: Evaluation of digital photogrammetric aerial based imaging systems – overview and results from the pilot centre. In: *IAPRSIS* XXXVIII (1-4-7/W5), on CD-ROM.

Dorninger, P., Nothegger, C., 2007. 3D segmentation of unstructured point clouds for building modelling. In: *IAPRSIS* XXXVI (3/W49A), pp. 191–196.

Drauschke, M., 2009. An irregular pyramid for multi-scale analysis of objects and their parts. In*: 7th IAPR-TC-15 WS on Graph-based Representations in Pattern Recogn.*, pp. 293-303.

Drauschke, M., Roscher, R., Läbe, T., Förstner, W., 2009. Improving image segmentation using multiple view analysis. In: *IAPRSIS* XXXVIII (3/W4), pp. 211-216.

Förstner, W., Wrobel, B., 2004. Robust estimation. In: Mc Glone (ed.), Manual of Photogrammetry. 5th ed., American Soc. Photogrammetry & R. S., Bethseda (Md), USA, pp. 103-110.

Haala, N., Hastedt, H., Wolff, K., Ressl, C., Baltrusch, S., 2010. Digital photogrammetric camera evaluation – generation of digital elevation models. *PFG* 2(2010) (in print).

Khoshelham, K., 2005. Region refinement and parametric reconstruction of building roofs by integration of image and height data. In: *IAPRSIS* XXXVI (3/W24), pp. 3–8.

Rottensteiner, F., Trinder, J., Clode, S., Kubik, K., 2004. Fusing airborne laser scanner data and aerial imagery for the automatic extraction of buildings in densely built-up areas. In: *IAPRSIS* XXXVI (B3), pp. 512-517.

Rottensteiner, F., Trinder, J., Clode, S., Kubik, K., 2005. Automated delineation of roof planes in LIDAR data. In: *IAPRSIS* XXXV (3/W19), pp. 221-226.

Song, Y., Shan, J., 2008. Building extraction from high resolution color imagery based on edge flow driven active contour and JSEG. In: *IAPRSIS* XXXVII (B3A), pp. 185-190.

Vosselman, G., Dijkman, S., 2001. 3D building model reconstruction from point clouds and ground plans. In: *IAPRS* XXXIV (3/W4), pp. 37–42.