REGION REFINEMENT AND PARAMETRIC RECONSTRUCTION OF BUILDING ROOFS BY INTEGRATION OF IMAGE AND HEIGHT DATA

Kourosh Khoshelham

Dept. of Surveying Engineering and Geomatics, University of Tehran, Iran
kourosh.kh@gmail.com

ABSTRACT:

The refinement of the features extracted from image data is a key issue in automated building extraction since feature extraction algorithms often result in incomplete features. This paper describes a method for the integration of image and Lidar height data, which leads to the refinement of initial image regions and the reconstruction of the parametric forms of roof planes. Region refinement is based on fitting planar surfaces to the height points that project into each image region. The number and parameters of the planar surfaces are used to split and/or merge the incomplete regions. Every refined region corresponds to a single plane in object space whose average height over the average terrain height determines whether it is a roof plane. Experiments with the proposed method demonstrate the capability of the method in region refinement and roof plane reconstruction.

1. INTRODUCTION

Image-based approaches to automated building extraction greatly rely on the completeness of the features extracted from the image data. Feature extraction algorithms, however, often result in incomplete features while many features are totally missed. The modification of feature extraction algorithms and the refinement of the extracted features, therefore, become key issues in automated building extraction. Image segmentation algorithms, in particular, are very likely to generate a partitioning of the image space that does not correspond to the partitioning of the object space by visible surfaces. This problem is generally referred to as over-segmentation and under-segmentation of the image data. As a consequence of the over-segmentation and under-segmentation problems, the automated system will fail to correctly reconstruct the 3D model of the building.

Various approaches to automated building extraction deal with incomplete and missed image features in different ways. Semi-automated approaches tend to focus the interactive part on image interpretation, which guarantees a reliable and complete feature extraction by a human operator (Gruen, 1998; Gruen and Wang, 1998). Perceptual relations between image features (Lowe, 1985; Wertheimer, 2001) have also been exploited in order to group incomplete low-level features into more complex high-level structures. Perceptual grouping methods (Boyer and Sarkar, 1999) have been widely used in automated building extraction for handling incomplete features (Dang et al., 1994; Lin et al., 1994; Bignone et al., 1996; Krisgnamachari, 1996; Henricsson, 1998; Jayness et al., 2003). These methods, however, concentrate on the grouping of image lines in most cases, and the problems in image regions have not been taken into account. Fuchs and Forstner (1995) address the over-segmentation and under-segmentation problems as inconsistencies in the relations between image features. Model-based approaches to automated building extraction (Fua and Hanson, 1988; Fischer et al., 1998; Gulch et al., 1999; Khoshelham and Li, 2004; Suveg and Vosselman, 2004) are less influenced by incomplete and missed features

due to the inherence of tight geometric constraints in the model. Nevertheless, the generation of correct model hypotheses in the model-based approaches depends on the completeness of the extracted image features.

While cues derived from image data, such as interrelations between features, may not be sufficient to infer the correct partitioning of the image into homogeneous regions, data from other sources can be very useful for this purpose. Laser scanner systems provide a direct measurement of the visible surfaces in object space; hence, height data from such a source have a great potential for the refinement of image regions. Integration of image and height data has been a topic of several previous works (Ameri and Fritsch, 2000; Rottensteiner et al., 2004); however, the use of height data for the refinement of extracted image features has not been brought into focus. This paper describes a method for the integration of image and Lidar height data, which leads to the refinement of initial image regions and the reconstruction of the parametric forms of roof planes.

The paper is structured in five sections. An examination of the over-segmentation and under-segmentation problems is given in the next section. Section 3 describes the method for the integration of image and height data, and the refinement of image regions. Experiments and results are shown in Section 4. Conclusions are made in Section 5.

2. CHARACTERISTICS OF SEGMENTATION PROBLEMS

The goal of segmentation algorithms is to partition an image into a number of homogenous regions that correspond to surfaces in object space. There are two problems common in all image segmentation algorithms:

- Over-segmentation: is the case where there exists a single surface in object space, but the algorithm partitions the

image of this surface into more than one region. The detected regions are called undergrown regions.

- Under-segmentation: is the case where there exist two or more surfaces in object space, but the algorithm detects only a single region in the image of the surfaces. The detected region is called an overgrown region.

The general approach to the correction of undergrown and overgrown image regions has been based on an ability of human brain that is often referred to as perceptual completion. In this approach, the incomplete features are completed so that the ideal interrelation is set up between them. For example, in terms of image edges and regions, the ideal relation is that there always must be an edge between two image regions. Therefore, if the edge does not exist, either the two regions are merged or the edge is added. Fuchs and Forstner (1995) identified various types on inconsistencies between image features that can be corrected by establishing the ideal interrelation (Fig. 1). There is, however, an ambiguity problem associated with perceptual completion. In many cases, the selection between the two possible choices is an ambiguous one. For example, in Fig. 2(C) there is no clue as to whether the edge must be removed, or the edge must be completed and the region split.
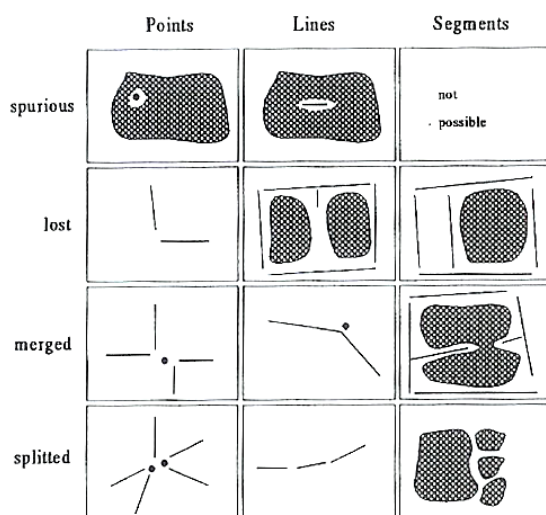

Fig. 1: Various types of inconsistencies in interrelations between different features (from Fuchs and Forstner, 1995).
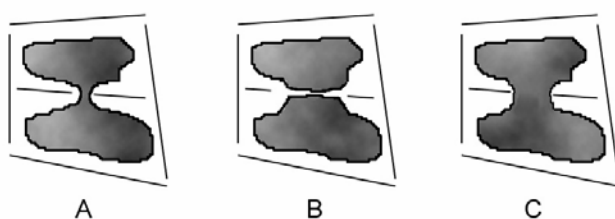

Fig. 2: Ambiguity problem in perceptual completion. A. The region is split because the edge is strong; B. The regions are merged because the edge is weak; C. An ambiguous case.

Another approach that has been used for the correction of undergrown and overgrown regions is the multi-resolution segmentation. In this approach, the image is resampled with a smoothing kernel into different resolution layers, and the segmentation algorithm is applied to each layer. Thus, undergrown regions may turn out merged in a smoother layer, and an overgrown region might be found correctly split in a sharper layer. The multi-resolution approach can be regarded as equivalent to selecting different smoothing parameters in the segmentation algorithm. Aside from the practical complications of selecting the right resolution layer or parameter setting in an automated fashion, this approach might as well fail due to intrinsic uncertainties present in image data. For example, when an edge between two surfaces in object space does not appear in the image due to low contrast, lighting condition or shadow the corresponding regions will overgrow in the segmented image regardless of the selected resolution or parameter setting. On the contrary, the presence of shadow or an undesirable object, e.g. an antenna, can lead to undergrown regions in all resolutions of the segmented image no matter what parameter settings are used.

The aforementioned problems with the segmentation of image data are less critical in height data. While surfaces in object space appear as homogenous regions in an intensity image, Lidar systems provide a direct measurement of these surfaces. In contrast, the edges of the surfaces are not measured accurately in Lidar height data as a result of the relatively low spatial resolution of the height points. Previous studies have shown that the segmentation of height data still faces the over-segmentation and under-segmentation problems due to measurement errors and the presence of undesirable objects (Hoover et al., 1996). These observations suggest that a combination of image and height data will result in a more complete segmentation of both sources. The next section describes the method for the refinement of image regions by integrating image and Lidar height data.

3. REGION REFINEMENT AND ROOF PLANE RECONSTRUCTION

In principle, the proposed method for region refinement is based on fitting planar surfaces to the height points that project into each image region. This is based on the assumption that buildings are planar objects; thus buildings with curved surfaces are not taken into account in this method. If the segmentation algorithm yields a correct partitioning of the image, every image region will have a single plane in object space that fits to its associated height points. In the case of under-segmentation, multiple planes will be found in the height points associated with the overgrown region. Similarly, in the case of over-segmentation, the planar surfaces found in neighbouring undergrown regions will be coplanar. Therefore, by examining the number and parameters of the planar surfaces found in the height data associated with every image region, it is possible to refine the initial segmentation by splitting overgrown regions and merging undergrown regions. This procedure can be regarded as a segmentation of height data guided by an initial segmentation of the image. An examination of the average height of the detected planes over the average height of the terrain provides evidence for finding roof planes. The reconstruction of the roof planes in this method is not influenced by the presence of vegetation, since the height data associated with vegetation regions are unlikely to fit in planar surfaces. The following sections describe the plane fitting, split-and-merge and roof determination processes.

3.1 Plane fitting process

A last echo Lidar DSM is used as the height data in this process. The height points from the DSM are projected into image regions, provided that orientation parameters are given. A robust regression method based on the least median of squares (LMS) is used for fitting planar surfaces to the height points in each image region. The height points that do not fit in the plane are treated as outlier points in the LMS estimator. Therefore, it is possible to detect multiple planes in the height data by iteratively applying the robust plane fitting algorithm to outlier points. The LMS estimator has a breakdown point of 50%; hence it can deal with half of the data points as outliers.

The detection of outliers in the LMS-type regression relies on the Random Sample Consensus (RANSAC) paradigm (Fischler and Bolles, 1981). The RANSAC algorithm is based on the selection of a number of sets of samples from the data (trial estimates). The number of the trial estimates is significantly reduced by specifying a confidence probability that at least one sample contains no outlier points. The plane parameters are calculated for every random sample (containing three data points), and the sample for which the median of the squared residuals (of all points) is minimum is selected as the best sample. The final plane is estimated using all inlier points, and is accepted if the standard deviation of its residuals does not exceed a maximum acceptable tolerance. The above steps are iteratively applied to outlier points until no new planes are found, or the number of the remaining points is not sufficient for plane fitting. Fig. 3 illustrates the process of plane fitting in an overgrown region.

3.2 Split-and-merge process

As mentioned above, in an ideal segmentation every image region is associated with a single plane in object space (assuming buildings are composed of planar faces). However, due to the over-segmentation and under-segmentation problems, in practice the region-plane correspondence does not exist, and multiple planes might be found in a single region, or multiple regions might be associated with a single plane. In the split-and-merge process, the results of the plane fitting process are used to detect overgrown and undergrown regions, and split and merge them respectively.

Overgrown regions are simply identified as regions in which multiple planes are found. Although only the detected planes are used in the reconstruction stage, image regions can as well be refined. An overgrown region is split by finding for each pixel the nearest height point and the plane that particular height point belongs to.

In order to merge undergrown regions it is necessary to establish adjacency relations between the image regions. The topology of image regions is most often represented in a graph structure where the graph nodes represent image regions, and arcs denote the adjacency relation between regions. Planar surfaces found in every pair of adjacent regions are examined to determine whether they are coplanar. The coplanarity check is based on the discrepancies between the slopes of planar

surfaces in adjacent regions, and the vertical distance between the two planes. If these values do not exceed a maximum tolerance, then the adjacent region-planes will be found coplanar and thus merged.

Every refined region is stored along with the parameters of its associated plane and the coordinates of the height points used for the computation of that plane. The coordinates of a height point include a height derived from the DSM and a height derived from the DTM. To determine the average height of a plane over the average terrain height the above two height values are subtracted and the results are averaged over the region. If the value of the average height difference is greater than a threshold, the plane is classified as a roof plane. A threshold of 3m to 4m is often convenient, since building roofs are unlikely to have a lower height.
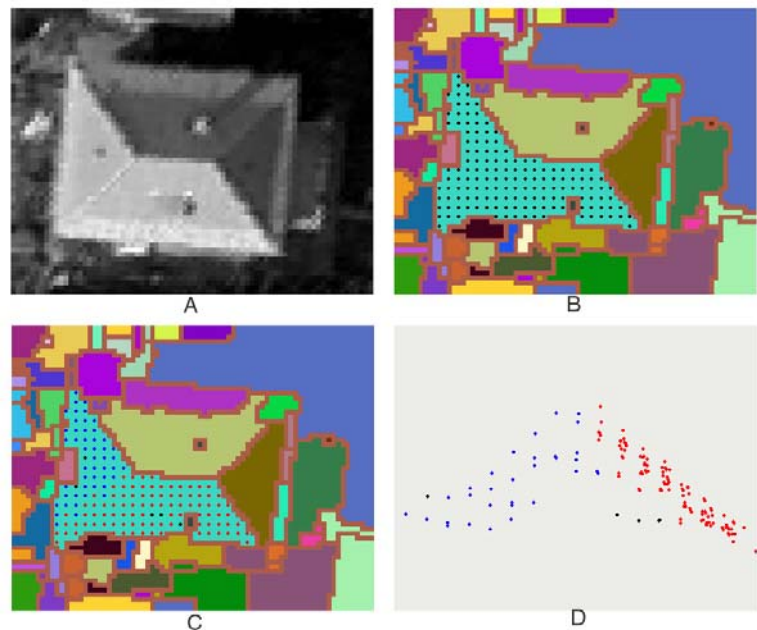

Fig. 3: The plane fitting process. A. An aerial image of a building; B. The initial segmentation of the image with the height points projected from the DSM to an overgrown region; C. Two planes are detected in the height points belonging to the overgrown region. Points in red belong to the first plane, points in blue belong to the second plane and points in black fit in neither of the planes; D. The height points in 3D view.

4. EXPERIMENTS

The method was tested with a set of Lidar data consisting of an orthoimage with 1m resolution, a last echo DSM and a DTM both with 0.5m resolution. The planimetric and altimetric accuracy of the DSM height points were 0.5m and 0.15m respectively. The location of the dataset is the city of Memmingen in southern Germany. Two cutouts were selected from the suburban area of the city. The images were segmented using a watershed algorithm with a smoothing parameter of 13.

Fig. 4 illustrates the process of region refinement and roof plane reconstruction for the Memmingen 1 cutout. As can be seen, the initial segmentation of the image (Fig. 4A) includes a significant number of overgrown and undergrown regions (Fig. 4C), which are refined by using the planar surfaces found

in height data (Fig. 4D). The application of the method to the Memmingen 2 cutout is shown in Fig. 5. Again, in most cases, the method is able to split overgrown regions, merge undergrown regions and separate roof planes from non-roof planes.

The performance of the method in terms of the completeness of the detected roof planes is summarized in Table 1. In total, out of 37 roof planes in the Memmingen 1 cutout, 32 planes are correctly detected. This rate is 23 out of 27 for the Memmingen 2 cutout. It is worth noting that planar surfaces that correspond to dormers and skylights are also considered as roof planes a long as their sizes are sufficiently large so that the projected height points form a plane. In other words, the level of detail for the reconstructed roofs is limited by the resolution of height data.

From the figures it can be observed that the performance of the method is not influenced by vegetation unless a roof plane is entirely covered by a tree canopy. But on the other hand, skylights and other small objects on the roofs that result in excessive over-segmentation of the image into very small regions have led to failure in a number of cases. There are also a few cases where the boundaries of a refined roof region do not conform to the boundaries of its associated plane. This is due to the inaccuracy of the height points on the boundaries of the roofs. These points have been detected as outliers, and thus left out in the plane fitting process. In general, the boundaries of the refined regions do not correctly show the boundaries of the roofs. Accurate extraction of roof boundaries are discussed elsewhere (Khoshelham, 2004).

The reconstructed roof planes were found to be within an acceptable range of accuracy. Since ground control points were not available for the location of the experiment, a number of checkpoints were derived from the DSM of the Memmingen 1 cutout in order to assess the accuracy of the reconstructed roofs. The altimetric accuracy of the reconstructed roof planes for the Memmingen 1 cutout was found to vary between 0.06m to 0.48m.

5. CONCLUSIONS

In this paper a method was presented for the refinement of image segmentation and reconstruction of parametric roof planes by integrating image and Lidar height data. The robustness of an image-based building reconstruction system greatly relies on the completeness of the features extracted from the images. Thus, by employing the region refinement process described in this paper one can expect a higher level of robustness in the reconstruction stage. Also the parametric roof planes reconstructed in this process can be effectively used in the final modelling step. In model-driven approaches to building reconstruction, the number and parameters of these roof planes can be used to guide the search for the right parametric models in a database of building models. In data-driven approaches, the method can be employed in conjunction with a boundary extraction algorithm to reconstruct the buildings with generic polyhedral models. In general, it can be concluded that integration of image and height data has a significant potential to reduce the complexities involved in building reconstruction, and improve the overall robustness of an automated building reconstruction system.

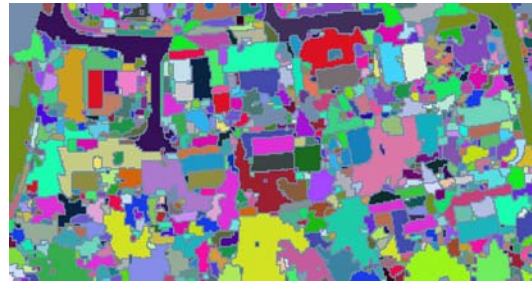Table 1: Completeness evaluation of the region refinement and roof reconstruction method.

	Memmingen 1	Memmingen 2
Total number of roof planes	37	27
Number of detected planes	33	24
Number of missed roof planes	5	4
Number of planes wrongly detected as roof	1	1

A

B

C

D

Fig. 4: The process of region refinement and roof plane reconstruction for Memmingen 1 cutout. A. Orthorectified aerial image of the scene; B. Last echo DSM of the scene; C. Initial segmentation of the image; D. refined regions and detected roofs.
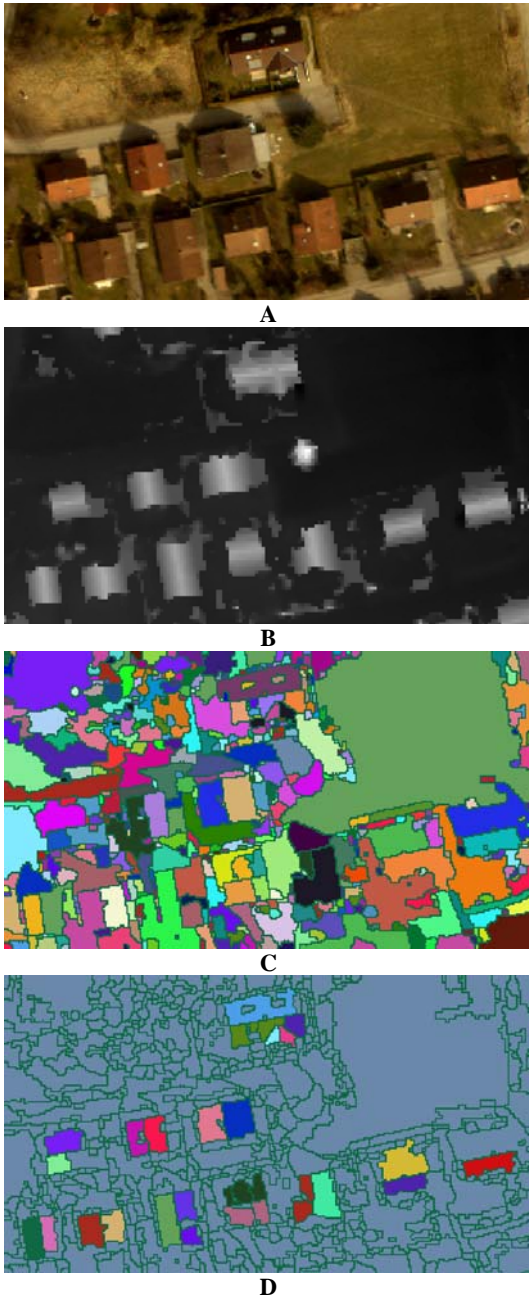
A

B

C

D

Fig. 5: The process of region refinement and roof plane reconstruction for Memmingen 2 cutout. A. Orthorectified aerial image of the scene; B. Last echo DSM of the scene; C. Initial segmentation of the image; D. refined regions and detected roofs.

ACKNOWLEDGEMENT

REFERENCES

Ameri, B. and Fritsch, D., 2000. Automatic 3D building reconstruction using plane roof structures, ASPRS Congress, Washington DC.

Bignone, F., Henricsson, O., Fua, P. and Stricker, M., 1996. Automatic extraction of generic house roofs from high resolution aerial imagery, ECCV '96: 4th European Conference on Computer Vision, Cambridge, UK, April 15-18, pp. 85-96.

Boyer, K.L. and Sarkar, S., 1999. Perceptual organization in computer vision: status, challenges, and potential. Computer Vision and Image Understanding, 76(1): 1-4.

Dang, T., Jamet, O. and Maitre, H., 1994. Applying perceptual grouping and surface models to the detection and stereo reconstruction of building in aerial imagery. In: H. Ebner, C. Heipke and K. Eder (Editors), ISPRS Symposium on Spatial Information from Digital Photogrammetry and Computer Vision, Munich, pp. 165-172.

Fischer, A., Kolbe, T.H., Lang, F., Cremers, A.B., Forstner, W., Plumer, L. and Steinhage, V., 1998. Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D. Computer Vision and Image Understanding, 72(2): 185-203.

Fischler, M.A. and Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24(6): 381-395.

Fua, P. and Hanson, A.J., 1988. Extracting generic shapes using model-driven optimization, Proceedings of the DARPA Image Understanding Workshop, Cambridge, Massachusetts, pp. 994-1004.

Fuchs, C. and Forstner, W., 1995. Polymorphic grouping for image segmentation, Fifth International Conference on Computer Vision, Massachusetts Institute of Technology, Cambridge, Massachusetts, pp. 175-182.

Gruen, A., 1998. TOBAGO -- a semi-automated approach for the generation of 3-D building models. ISPRS Journal of Photogrammetry and Remote Sensing, 53(2): 108-118.

Gruen, A. and Wang, X., 1998. CC-Modeler: a topology generator for 3-D city models. ISPRS Journal of Photogrammetry and Remote Sensing, 53(5): 286-295.

Gulch, E., Muller, H. and Labe, T., 1999. Integration of automatic processes into semi-automatic building extraction, ISPRS Conference, Automatic Extraction of GIS Objects from Digital Imagery, Munich, Germany.

Henricsson, O., 1998. The Role of Color Attributes and Similarity Grouping in 3-D Building Reconstruction. Computer Vision and Image Understanding, 72(2): 163-184.

Hoover, A., Jean-Baptiste, G., Jiang, X., Flynn, P.J., Bunke, H., Goldgof, D.B., Bowyer, K., Eggert, D.W., Fitzgibbon, A. and Fisher, R.B., 1996. An experimental comparison of range image segmentation algorithms. IEEE transactions on pattern analysis and machine intelligence, 18(7): 673-689.

Jaynes, C., Riseman, E. and Hanson, A., 2003. Recognition and reconstruction of buildings from multiple aerial images. Computer Vision and Image Understanding, 90(1): 68-98.

Khoshelham, K. and Li, Z.L., 2004. A model-based approach to semi-automated reconstruction of buildings from aerial images. Photogrammetric Record, 19(108): 342-359.

Krisgnamachari, S. and Chellapa, R., 1996. Delineating buildings by grouping lines with mrfs. Transactions on Image Processing, 5: 164-168.

Lin, C., Huertas, A. and Nevatia, R., 1994. Detection of building using perceptual grouping and shadows, Conference on Computer Vision and Pattern Recognition, Seattle, Washington, pp. 62-69.

Lowe, D.G., 1985. Perceptual organization and visual recognition. Kluwer Academic Publishers, Hingham, MA, 162 pp.

Rottensteiner, F., Trinder, J., Clode, S. and Kubik, K., 2004. Fusing airborne laser scanner data and aerial imagery for the automatic extraction of buildings in densely built-up areas. International Archives of Photogrammetry and Remote Sensing, 35(B3): 512-517.

Suveg, I. and Vosselman, G., 2004. Reconstruction of 3D building models from aerial images and maps. ISPRS Journal of Photogrammetry and Remote Sensing, 58: 202-224.

Wertheimer, M., 2001. Laws of organization in perceptual forms. In: S. Yantis (Editor), Visual perception. Key readings in cognition. Psychology press, pp. 431.