AUTOMATIC BUILDING OUTLINING FROM MULTI-VIEW OBLIQUE IMAGES

J. Xiao

Faculty of Geo-Information Science Observation, University of Twente, 7514AE Enschede, The Netherlands -
jing@itc.nl

Commission III, WG III/4

KEY WORDS: Building outlining, oblique images, AdaBoost, region growing, image segmentation, stereo matching

ABSTRACT:

Automatic building detection plays an important role in many applications. Multiple overlapped airborne images as well as lidar point clouds are among the most popular data sources used for this purpose. Multi-view overlapped oblique images bear both height and colour information, and additionally we explicitly have access to the vertical extent of objects, therefore we explore the usability of this data source solely to detect and outline buildings in this paper. The outline can then be used for further 3D modelling. In the previous work, building hypotheses are generated using a box model based on detected façades from four directions. In each viewing direction, façade edges extracted from images and height information by stereo matching from an image pair is used for the façade detection. Given that many façades were missing due to occlusion or lack of texture whilst building roofs can be viewed in most images, this work mainly focuses on improve the building box outline by adding roof information. Stereo matched point cloud generated from oblique images are combined with the features from images. Initial roof patches are located in the point cloud. Then AdaBoost is used to integrate geometric and radiometric attributes extracted from oblique image on grid pixel level with the aim to refine the roof area. Generalized contours of the roof pixels are taken as building outlines. The preliminary test has been done by training with five buildings and testing around sixty building clusters. The proposed method performs well concerning covering the irregular roofs as well as improve the sides location of slope roof buildings. Outline result comparing with cadastral map shows almost all above 70% *completeness* and *correctness* in an area-based assessment, as well as 20% to 40% improvement in *correctness* with respect to our previous work.

1. INTRODUCTION

Automatic building detection is important for many applications, for instance map updating, city modelling and urban planning. Various data sources have been used for building detection, including airborne images (Müller and Zaum, 2005; Karantzalos and Paragios, 2009), height data in form of DSM or DEM (Ma, 2005; Lu et al., 2006), or laser scanning data (Kim and Shan, 2011). However, in order to overcome the limitation from either data set, researches on integration of them are quite active nowadays (Awrangjeb et al., 2010; Kabolizade et al., 2010; Khoshelham et al., 2010).

Oblique airborne images, taken by multiple cameras covering different view directions, are currently available in several cities and are being systematically captured, e.g. by Pictometry[1] or Slagboom en Peeters B.V[2]. Having a large tilt angle as the unique feature, oblique images depict building façade structures. This unique feature of oblique images is useful for building detection and verification (Nyaruhuma et al., 2010). These images have overlap, and thus enable the generation of 3D points as discussed in (Xiao et al., 2012).

The method presented in this paper aims to automatically detect and map building outlines from oblique images. The outlines can be used for updating cadastral maps or further 3D modelling. To widen its application this method is completely independent from any other data source. The paper is based on the previous work on building façade detection from multi-view oblique images (Xiao et al., 2012). A simple box model was used in the previous work to indicate the existence of a detected building. Little effort was made on completing the building area, especially on large irregular buildings. Therefore, this paper aims at starting from an inaccurate box model to complete the building area by adding roof information.

Similar as the fusion of lidar data and imagery, this method integrates geometric and radiometric information of the images. The former is extracted from a point cloud generated by multi-view stereo matching while the latter is used in form of image segmentation and line extraction. AdaBoost (adaptive boosting) is implemented as a means to combine several features for roof classification. Assuming that the requirement on more prior knowledge might have more limits on the application of this method, we attempt to look for independent features that can be trained very locally but be used globally.

This paper begins with the previous work on building detection and shares the same data. A brief description on the image data and the method on generating building hypotheses will be given first. The result of the method described in this paper is compared to the previous result. Large scale cadastral map with 10 cm to 20 cm accuracy and lidar data are used as reference.

2. REVIEW OF PREVIOUS WORK AND DATA USED

In this section, a brief description of the method on building detection from oblique airborne images is presented, including results. For details see (Xiao et al., 2012).

---

[1]
  http://www.pictometry.com/index.php?option=com_content&view=article&id=76&Itemid=85

[2] http://www.slagboomenpeeters.com/obliek.html

2.1 Multi-view oblique images

The used data is composed of images viewing from four directions. For one point in object space, four to eight images from different perspectives are available. The images have a tilt angle around 50 degrees, and a resolution of 10 to 16 cm. All the images used in the experiment were oriented using the method described in (Gerke, 2011). The RMSE at check points in object space was around 20 cm for all three components after the self-calibration bundle adjustment. We keep using the same data sets so as to reuse the results as well as to check the improvement on the building outlines.

2.2 Façades detection

Building façades are detected in each viewing direction separately. A pair of images is used to generate one 3D façade. The detection employs extracted façade structures from images and height gradients from a projected 3D point cloud generated from that image pair.

2.2.1 Façade structures: façade structures are retrieved from our assumption about vertical and horizontal structures being present at building façades. In one oblique image, vertical and horizontal structures present as almost plumb and parallel lines respectively. Therefore, façade structures are extracted via line extraction in a single image, following by locating places where there are dense plumb and parallel lines.

2.2.2 Height gradients: height gradients at the façade plane are supposed to be larger than on the ground or on the roof plane. The calculation of height gradients is based on the 3D point cloud from the image pair. To achieve the integration with the former evidence, heights of the 3D points are projected into the source image pair and interpolated to a height map. The height gradient is calculated using this height map.

2.2.3 Façade patch detection: this operation is based on the multiplication of the two described features. Possible façade patch pixels are selected on setting a threshold to the multiplied value. Experiments showed that the detection results are not very sensitive to changes in the threshold.

2.2.4 Façade patch generation in 3D: Back projecting image pixels in a façade patch with height from height map results as a cluster of points in 3D. A vertical plane is fit to the points to reconstruct the façade plane. This process is done parallel in the image pair, therefore the two façade plane sets can be used to verify each other.

2.3 Building hypotheses generation in 3D space

Building hypotheses are generated by employing a box model to combine façade hypotheses from four directions. One box hypotheses is constructed with at least one detected façade on its sides, and the best case is that its four sides are all supported. For the unsupported side(s), some simple assumptions (e.g. width) are made to complete the box.

2.4 Results of the described method and problems

In a test area with 333 buildings, more than 85% of buildings are successfully detected, whilst in the area where complete image pairs from four directions available, the percentage can reach 95%.

The main deficiency of the method is caused by the box model. It is incapable to delineate irregular buildings. Therefore a better outlining strategy needs to be developed.

3. OVERVIEW OF THE PROPOSED WORK ON BUILDING OUTLINING

Starting with building box hypotheses from previous work, this work aims at outlining the detected buildings. The previous work mainly based on building façades, but many façades are missing due to occlusion or lack of texture. Since roofs have less occlusion, they are involved in this work to complete the whole building for outlining. We maintain the idea of integrating point cloud with image information, for the main reason that although the point cloud is able to provide relatively accurate 3D position, the matching method doesn't perform well on edges or non-textured areas. This insufficiency is expected to be compensated by radiometric information on the images. One of the advantages of taking multi-view images is that even when the contrast on one edge is not significant in some images, it may be better in others.

The proposed method mainly consists out of two stages: initial roof locating (Section 4) and roof area growing (Section 5). Since the location and height of the building hypotheses is not precise, causing difficulties to project them into images, we start with initializing roof plane(s) from the generated point cloud. The off-ground point patches from surface growing are classified into roof and façade based on their heights and normal directions.

A growing approach (refer to Section 5.4) is developed in this paper to complete the whole roof area. It makes used of features from extracted roof points (Section 5.2) and features from images (Section 5.3). In order to integrate the point cloud with multi-view images, a 2.5D representation of the scene is being used. The area is gridded and each cell has the height from the point cloud as an attribute. Since the height of each quadratic cell is known, the features from images are extracted by projecting the cell into the images. Building outlining (Section 5.5) is based on the result from roof growing, but also taking extracted wall and slope roof direction as restrictions.

4. INITIAL ROOF LOCATION FROM POINT CLOUD

Taking one building hypothesis at a time, a point cloud is generated for it. Roof patches and façade patches are extracted from the point cloud to be used as the growing seeds and constraints for outlining respectively for later steps.

4.1 Point cloud generation

Based on the input box model, a region of interest (ROI) is located for each building. Then images are selected to generate point cloud specifically in the ROI. This helps to maximize the number of points for the target building.

Point clouds are generated using the patch-based multi-view stereo (PMVS) matching method described in (Furukawa and Ponce, 2010). A patch indicates a small oriented rectangle in the method. Features are first matched across multiple images to generate sparse patches, and then expand to its neighbours to

make patches dense. A filtering step is used to remove outliers and to avoid matching within occluded areas.

This PMVS approach generates acceptable accurate point clouds covering the roof, except in textureless and occluded areas. Because of this, the point cloud may not be as regular as for instance from lidar; especially in homogenous roof areas, points may be missing. An alternative could be to apply regularization in order to densify the point cloud, like in global approaches like SGM (Semi-global matching, (Hirschmüller, 2008)). However, since in the next step we merely fit planes to the point cloud, those missing areas will probably not have significant effects on the result. Another reason not to use regularization is that those approaches can also be seen as another kind of model-driven interpolation: If there is no texture, the gaps are bridged using smoothness constraints, thus might bring in some additional uncertainty.

4.2 Roof and façade patches composition

A surface growing process is applied, using the point cloud. Vertical patches are firstly extracted as façade patches. Then points below a certain height from the lower points of the vertical patches are removed as ground points.

Patches above the façade patches are grouped according to their plane normal and locations. Horizontal planes are formed from respective patches if they maintain a certain distance and height difference threshold. Nearby slopes patches are grouped when they have a similar normal direction. Symmetric patch groups are again tested on the position of their intersection and their relative location to check whether they belong to the same gable roof building. Although symmetric planes are still grown separately in the later step, this helps to reduce inaccuracy caused by the noise of the point cloud, especially for the roof parts where only a few points are generated.

Sloped roofs are usually fragmented due to small structures like dormers and noise in the point cloud. For each of those fragments it is tested whether it complies with one of the surrounding larger roof patches and is added to it if applicable. Lastly, small patches not belonging to any group or small groups with few points are deleted. An example of roof patch segmentation is shown in Fig. 2.
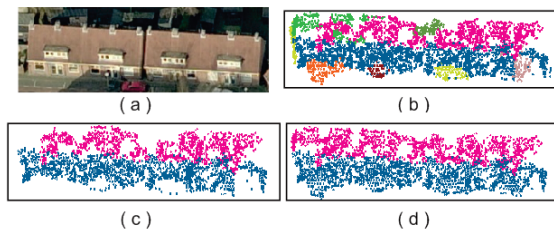

Figure 1 Initial roof and façade patches from point cloud. (a) sample building in oblique images; (b) initial patches above façade patches after surface growing; (c) grouped roof patches; (d) expanded roof patches. In (c) and (d), one colour indicates one group. Image © Blom

5. OUTLINING BASED ON ROOF GROWING WITH IMAGE INFORMATION

5.1 General principle of roof growing

Results of the initial roof patches are considered as reliable, but they are usually too sparse and incomplete for the outlining.

Therefore, a growing process is developed to expand the initial roof patch to reach edges.

After processing of the geometric information from the point cloud, the aim of this stage is to integrate radiometric information from images to expand the roof points. Assuming that image segmentation intends to preserve roof regions that have homogenous colour, and leave the roof edge between the segments, images are segmented to guide the growing process.

For this process, a 2.5D representation of the scene is used. The area is divided into a regular grid. Each cell is treated as a carrier of the features from both object space and images space. The growing is based on a binary classification on the neighbourhood of certified roof points in the gridded area. AdaBoost is used to perform the classification into roof and non-roof cell. The implemented strategy constitutes a growing algorithm, because seed cells are selected and classified according to some criteria. After each classification, features for other, yet unclassified cells are updated.

In order to reduce the training effort, features which can easily vary from building to building, such as colour, are avoided. Selected features are sketched in Figure 2 and elaborated in the following two sections.
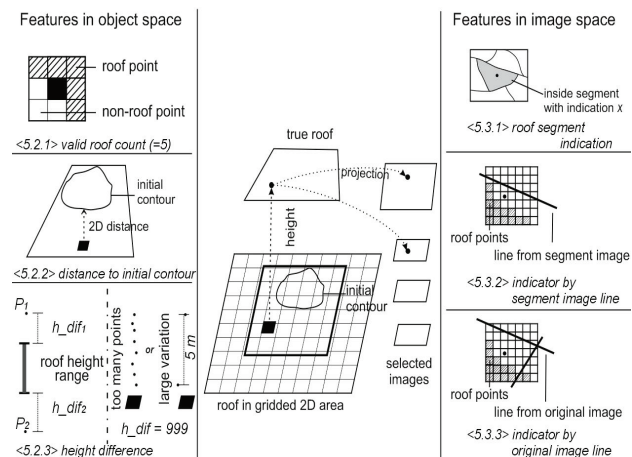

Figure 2. Selected features from both object space and image space used for roof growing

5.2 Features in object space

Features from object space are obtained by interpolating the point cloud into the ortho-projected cells.

5.2.1 Valid roof count in the neighbourhood: this feature encodes the number of already certified roof cells within a window. This is to exploit the assumption that it is more likely to have a roof cell if adjacent cells are also labelled as roof (and vice versa).

5.2.2 2D Distance to initial contour: The contour of the certified roof patches is used as the initial contour. Assuming that the further a grown point from the certified roof patches, the less support it gets, the distance from the initial contour is taken as a feature (0 if inside the contour). The distance is measured in 2D to be compatible with various slopes. This feature is becoming more important when the growing is advancing from the roof centre towards edges.

5.2.3   Height difference from roof: The height range for a specific roof is calculated from the initial certified roof patches. A roof point should not have height far outside this range. The average of all the points projected into a cell is taken as the cell height, except two cases listed below indicating the cells on the façade. The absolute difference from the cell height to the height range is calculated as the height difference (0 if within the height range).

5.3   Features in each select image

Several images are selected for image features. Multiple images from different perspectives may alleviate the effects of low contrast at roof edge from one direction. The height of a cell is firstly calculated from the initial roof plane. The point is then projected onto images. Since it is assumed that the initial roof contour covers the actual roof and only roof cells are considered in the following, problems from re-projecting cells at a wrong heights cannot be expected. The average of the roof segment indication values from all selected images is taken as the final value. The minimum value from all images is used for the latter two respectively to overcome the bad effects of the low contrast from a certain perspective.

5.3.1   Roof segment indication: A graph-based method is used to segment individual images (Felzenszwalb and Huttenlocher, 2004). The segment indication is a function of two variables. It is normalized to [0, 1].

➢   overlap with the initial roof patch contour: the higher percentage of the segment covered by the roof contour, the larger this indicator value;
➢   number of the façade points: 3D wall points are projected onto the images to check whether the segment or part of the segment is a façade. The count of the number lowers the roof segment indicator.

5.3.2   Indicator counting the on/off-roof pixels on the sides of extracted lines from segment image: long straight lines extracted (Förstner, 1994) from the segment image give a strong evidence for the existence of edges, which represent roof boundaries if around roof segments. To ensure that a growing pixel does not exceed the roof boundary, a searching distance is defined to look for extracted lines around the projected grid point on the image. It is expected that a target pixel should have many on-roof pixels on the same side of an extracted line versus few on the other side, in contrast with off-roof points. An n × n window is placed on the projection of the target pixel onto the edge to count the evidence. See Figure 3 for a sketch.
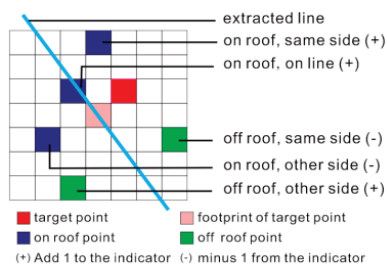

Figure 3. Calculation of the indicator from extracted line

In the ideal case to classify a target pixel to be on-roof, it should have all on-roof pixels on the same side with the extracted line, but off-roof pixels on the other side. In this case, the value will reach the maximum (n × n). This maximum is also assigned to pixels which cannot find any extracted lines in the search distance, meaning that they don't have any roof bound constraint. If multiple lines are found, the indicator takes the minimum value from all lines in order to increase its risk to be off-roof.

5.3.3   Indicator counting the on/off-roof pixels on the sides of extracted lines from original image: This feature is similar to the former one. The two features are separated due to the reason that the extracted lines here solely rely on strong edges in the image, while the one retrieved from segments is less dependent on strong edges. Lines from original images can be affected by the unimportant roof structures or other textures in the image, while those segment edges can be arbitrarily, and have less semantic meaning. The use of both features (5.3.2 and 5.3.3) intends to put higher weights on the actual roof boundary.

5.4   Growing approach

5.4.1   Seed selection and growing sequence: First, the valid roof count (feature of 5.2.1) is calculated for all points from the initial roof patch and sorted in descending order. Cells completely surrounded by roof points are omitted because no growing is needed for them. Then we start growing from the ones having the maximum count for this particular feature. Through the growing procedure, the valid roof count will be updated.

A queue is built to manage the growing sequence. The valid seeds with maximum neighbourhood count are initially added to the queue. After processing the first seed in the queue its unknown neighbours are added at the end of the growing queue, given that the seed is classified as roof cell. The growing ends when all the seeds in the queue are finished.

5.4.2   Training AdaBoost classifier: Several sample buildings with interactively extracted outlines are used for training. In general, the points only have two classes: *on-roof* and *off-roof*. Starting with the points from initial roof patches, unknown points around them are added in the queue for further growing. Each unknown point participating in the growing procedure is treated as one training sample. If the sample is inside the reference outline, its class label is set to *on-roof*, otherwise *off-roof*. By this means the growing procedure is also applied for training, which ensures that the dependencies between the individual features, especially in the transition area around the contour are considered also during training.

5.5   Outlining

Building outlining is mainly based on the result from roof growing. Restrictions, such as extracted façade patches and slope roof directions are used to improve the outline. Flat and slope roof buildings are treated separately in this step.

5.5.1   Outlining of flat roof buildings: a rough outline is firstly derived from the grown points and generalized by Douglas-Peuker algorithm. Extracted façade patches from Section 4.2 are selected to intersect the flat plane to adjust the outline. The rest part of the outline is then improved by checking parallelism, perpendicularity and collinearity with the edges from façades.

5.5.2   Outlining of slope roof buildings: rectangles are used to fit slope roof planes. Direction of the upper and lower bound of a rectangle are fixed to be perpendicular with the roof normal. The sides are closed according to the extension of the grown roof points. If two symmetric slopes planes are available for a building, the side extensions are adjusted to their average.

6. EXPERIMENTS AND RESULTS

6.1 Experiment design

This preliminary experiment was designed to test the improvement on delineating building outlines by the proposed method compared to the previous result in (Xiao et al., 2012). Previous results showed that buildings cannot be separated in the images if the gap between them is smaller than 3 m. Those nearby buildings were clustered beforehand. In total 50 sloped roof building clusters and eight irregular shaped buildings were selected from the previous test data set. To check whether the classifier in roof growing step can be trained efficiently, only two irregular flat roof buildings plus three slope roof buildings are selected for training.

The result after roof growing was first compared with the 3D lidar data to roughly check whether the 3D geometry of grown roof points was correct. The accuracy of the 2D building outlines was tested using a cadastral map, while result was also compared to the previous result to check the improvement. Area-based assessment used in (Xiao et al., 2012) was also used to check the 2D accuracy. For each cluster of buildings it calculated the intersection of the real building area and the outlined area (*TP*). The area in the real building but not in outlined area denoted *FN,* while the controversy denoted *FP*. Then *completeness* and *correctness* are defined as *TP/(TP+FN)×100%* and *TP/(TP+FP)×100%* respectively.

6.2 Result in 2.5D

Examples of the growing results are shown in Figure 4. The first column of Figure 4 shows the building hypotheses as resulting from the previous approach (section 2), overlaid onto an ortho image as well as the final outline. The second column shows the initial point matching result from PMVS, while the third column depicts the grown roof points, i.e. after classification. The height is coded as a colour in all height images. As a reference, the fourth column shows lidar data of the same area for the comparison with result from images.

Because the major difficulty in the previous result is the delineation of irregular roof buildings, three of them are selected (Figure 4 (a) - (c)) vs. two slope roof buildings ((d) & (e)). Previously irregular roof buildings were outlined by multiple boxes, and left much uncovered area. After the application of the new method, they are almost completely covered. The big hole in the initial point cloud on the flat roof plane in building (b) is filled in by the growing step. The missing part to the middle right in building (c) from the initial roof patch is also successfully recovered. However the growing fails to exactly recover the part in the lower right corner.

Tests are also made on slope roof buildings. Because of the constraint from detected façades, the previous result preformed fine on the location of front and back façades, but had problems on the sides. In Figure 4 (d) & (e), the initial point had holes in the roof plane as well as on the edges, and the grown process was able to fill the holes. But the growing method can only work with initially detected planes from the initial point cloud. Figure 4 (d) is an example of a cluster of nearby buildings. We took them as one in the following 2D assessment.

6.3 Result in 2D

During the roof growing process, four out of 53 building clusters missed. Three of them are individual small houses, so that the method failed to recognize them from small number of points from point cloud. The fourth one is a complex slope roof building (Figure 5(a)). It has small slope planes with diverse directions on the roof which could not be distinguished from noises in the process.

Successfully outlined buildings were compared with cadastral map and the previous result. Seeing from some examples shown in Figure 5, the delineation of the sides of slope roof buildings are much improved. The proposed method also succeeded on outlining irregular and regular flat roof buildings (Figure 5 (b) & (c) respectively).
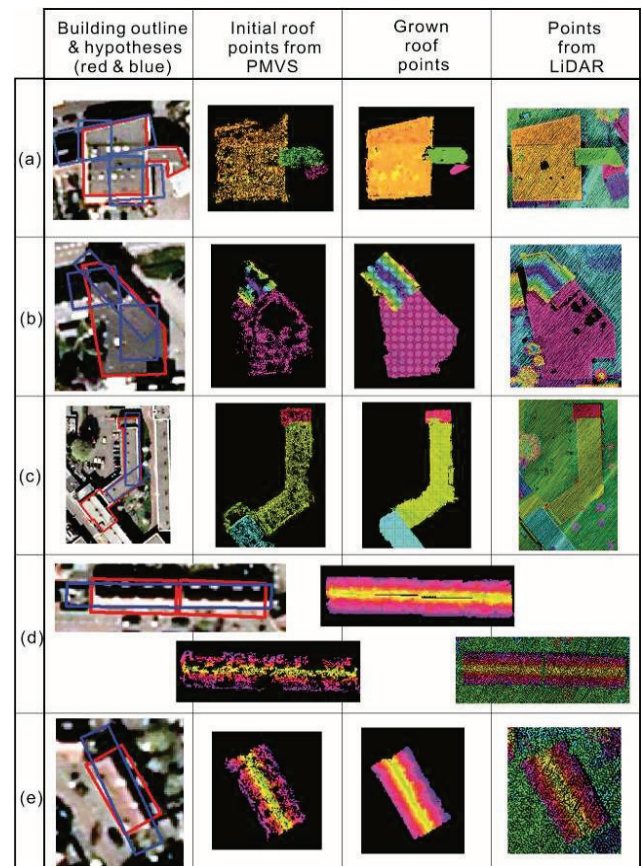

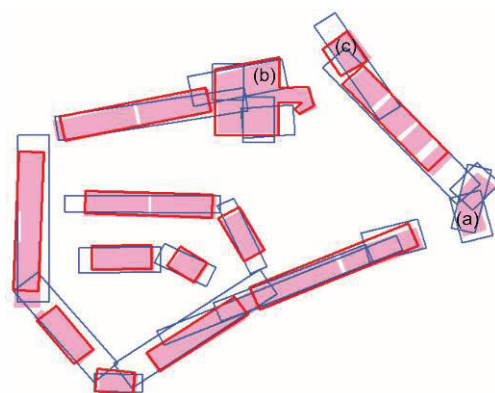
Figure 4. Result of roof growing approach

Figure 5. Comparison of outline (red line) with cadastral map (solid pink) and previous box model (blue line)

Figure 6 shows the overall result from the area-based assessment descried in Section 6.1. Irregular roof buildings are separated from slope roof buildings. Figure 6 (a) presented the

completeness and *correctness* referencing the cadastral map, while the comparison over the same test data is presented in (b). The area-based *completeness* and *correctness* of almost all tested buildings have reached 70%, many above 90%. The main inaccuracy for slope roof buildings happed at the ends of a row of nearby houses and the gaps in between. Height jumps is the main reason for incompleteness in irregular roof buildings.

The improvement is calculated by subtracting the previous *completeness*/*correctness* from the one in current result. The current *completeness* is in general even a bit lower, because the initial box models usually outbound the real buildings. For the same reason the current approach outperforms the former one in terms of *correctness* by 20% to 40%.
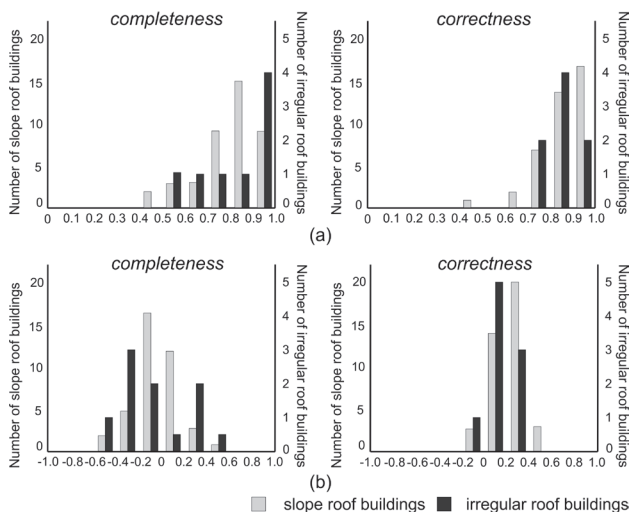

Figure 6. Area based assessment (a) compared with cadastral map and (b) improvement from previous result

7. CONCLUSION AND FUTURE WORK

This paper presents a complete approach of automatic building detection from multi-view oblique images solely. With the assistance of a point cloud derived by image matching, buildings can be delineated, starting from an initial box model. The fusion of the generated point cloud with image features alleviates the impact of noise and sparseness of the point cloud onto the final result. The selected features are independent from individual buildings, therefore the training result from a small dataset can be used for larger areas. Another advantage of this work is that it has no limitation on building shape, although it may face difficulties when a building roof is made up by many small roof faces of different directions.

This method is still incapable in dealing with complex slope roof buildings since the patch size is one of the criteria the method used to distinguishes roof patches from noise. Some preprocessing removing the noise on the initial generated point cloud might be helpful in this case.

REFERENCES

Awrangjeb, M., Ravanbakhsh, M. and Fraser, C.S., 2010. Automatic detection of residential buildings using LIDAR data and multispectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(5), pp. 457-467.

Felzenszwalb, P.F. and Huttenlocher, D.P., 2004. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2), pp. 167-181.

Förstner, W., 1994. A framework for low level feature extraction. In: *Computer Vision (ECCV'94)*, Stockholm, Sweden, Vol., pp. 383-394.

Furukawa, Y. and Ponce, J., 2010. Accurate, dense, and robust multi-view stereopsis. *IEEE transactions on pattern analysis and machine intelligence*, 32(8), pp. 1362-1376.

Gerke, M., 2011. Using horizontal and vertical building structure to constrain indirect sensor orientation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), pp. 307-316.

Hirschmüller, H., 2008. Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2), pp. 328-341.

Kabolizade, M., Ebadi, H. and Ahmadi, S., 2010. An improved snake model for automatic extraction of buildings from urban aerial images and LiDAR data. *Computers, Environment and Urban Systems*, 34(5), pp. 435-441.

Karantzalos, K. and Paragios, N., 2009. Recognition-driven two-dimensional competing priors toward automatic and accurate building detection. *IEEE Transactions on Geoscience and Remote Sensing*, 47(1), pp. 133-144.

Khoshelham, K., Nardinocchi, C., Frontoni, E., Mancini, A. and Zingaretti, P., 2010. Performance evaluation of automated approaches to building detection in multi-source aerial data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), pp. 123-133.

Kim, K. and Shan, J., 2011. Building roof modeling from airborne laser scanning data based on level set approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4), pp. 484-497.

Lu, Y.H., Trinder, J.C. and Kubik, K., 2006. Automatic building detection using the Dempster-Shafer algorithm. *Photogrammetric Engineering & Remote Sensing*, 72(4), pp. 395-403.

Ma, R.J., 2005. DEM generation and building detection from Lidar data. *Photogrammetric Engineering & Remote Sensing*, 71(7), pp. 847-854.

Müller, S. and Zaum, D.W., 2005. Robust building betection in aerial images. In: *Proceedings of the ISPRS Workshop CMRT 2005*, Vienna, Austria, Vol. 34, pp. 143-148.

Nyaruhuma, A.P., Gerke, M. and Vosselman, G., 2010. Evidence of walls in oblique images for automatic verification of buildings. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, Paris, France, Vol. 38, pp. 263-268.

Xiao, J., Gerke, M. and Vosselman, G., 2012. Building extraction from oblique airborne imagery based on robust façade detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 68, pp. 56-68.