

REFINING BUILDING FACADE MODELS WITH IMAGES

Shi Pu and George Vosselman

International Institute for Geo-Information Science and Earth Observation (ITC)
Hengelosestraat 99, P.O. Box 6, 7500 AA Enschede, The Netherlands
spu@itc.nl, vosselman@itc.nl

Commission III/4

KEY WORDS: building reconstruction, data fusion, image interpretation

ABSTRACT:

Laser data and optical data have a complementary nature to 3D features' extraction. Building reconstruction by fusion of the two data sources can reduce the complexity of approaches from either side. In this paper we present a model refinement method, which uses the strong lines extracted from close-range images to improve building models reconstructed from terrestrial laser point clouds. First, model edges are projected from model space to image space. Then significant line features are extracted from an image with Canny edge detector and Hough transformation. Each model edge is then compared with its neighboring image lines to determine the best match. Finally the model edges are updated according to their corresponding image lines. The refinement process not only fixes certain geometry errors of the original models, but also adapts the models to the image data, so that more accurate texturing is achieved.

1 INTRODUCTION

The technique of automated building facade reconstruction is useful to various applications. For urban planning, building facade models provide important references to the city scenes from the street level. For historical building documentation, a large number of valuable structures are contained on the facades, which should be recorded and reconstructed. For all virtual reality applications with users' view on the street, such as virtual tourism and computer games, the accuracy and/or realistic level of the building facade models are vital to successfully simulate an urban environment.

A number of approaches (Dick et al., 2001, Schindler and Bauer, 2003, Frueh et al., 2005, Pollefeys et al., 2008) are available for reconstructing building facades automatically or semi-automatically. Close range image and terrestrial laser point cloud are the commonly used input data. Image based building reconstruction has been researched for years. From multiple 2D images captured from different positions, 3D coordinates of the image features (lines for example) can be calculated. Although acquisition of images is cheap and easy, the difficulties of image understanding make it still difficult to automate the reconstruction using only images. Laser altimetry has been used more and more in recent years for automated building reconstruction. This can be explained by the explicit and accurate 3D information provided by laser point clouds. Researches (Vosselman, 2002, Frueh et al., 2004, Brenner, 2005) suggest that the laser data and images are complementary to each other, and efficient integration of the two data types will lead to a more accurate and reliable extraction of three dimensional features.

In the previous work we presented a knowledge based building facade reconstruction approach, which extracts semantic facade features from terrestrial laser point clouds and combines the feature polygons to water-tight polyhedron models (Pu and Vosselman, 2009). Some modeling errors still exist, and some of them can hardly be corrected by further exploiting the laser data. In this paper, we present a model refinement method which uses strong line features extracted from images to improve the building facade models generated from only terrestrial laser points. The refinement not only fixes the models' geometry errors, but

also solves inconsistencies between laser and image data, so that a more accurate texturing can be achieved.

This paper is organized as follows. Section 2 gives an overview of the presented method. Section 3 provides the context research of building reconstruction from terrestrial laser scanning. Section 4 explains the preprocessing steps such as perspective conversion and spatial resection, to make images usable for refining models. Section 5 elaborates the image processing algorithms used for significant line extraction and the matching and refinement strategies. Experiments on three test cases are discussed in section 6. Some conclusions and outlooks are drawn in the final section.

2 METHOD OVERVIEW

A building facade model may contain various errors. For a model reconstructed from terrestrial laser points, the model edges may have certain offset with their actual positions. These errors are caused by gaps in laser points and the limitations of laser data based reconstruction algorithms. Edges are delineated accurately in images. After registering to the model space, image lines can provide excellent reference from which the model edge errors can be fixed. Another necessity of this refinement is to solve the inconsistencies between the laser space and the image space, so that accurate texturing can be achieved.

Before starting the refinement, a 2D image needs to be referenced to the 3D model space, a problem often referred as spatial resection in photogrammetry. We use the standard resection solution of collinearity equations, which requires minimum three image points with their coordinates in model space. To find significant line features from an image, we first detect edges using the Canny algorithm (Canny, 1986), then apply the Hough transform to further extract strong line features from edges. Then model edges are projected to the image space and matched with the image lines. The best match is determined by the geometric properties of candidates and the geometric relations between candidates and the model edge. Finally each model edge with successful matching is projected to the matched image line accordingly, and model edges without any matching are also adjusted to maintain a well shape. Figure 1 gives a flowchart of the refinement process.

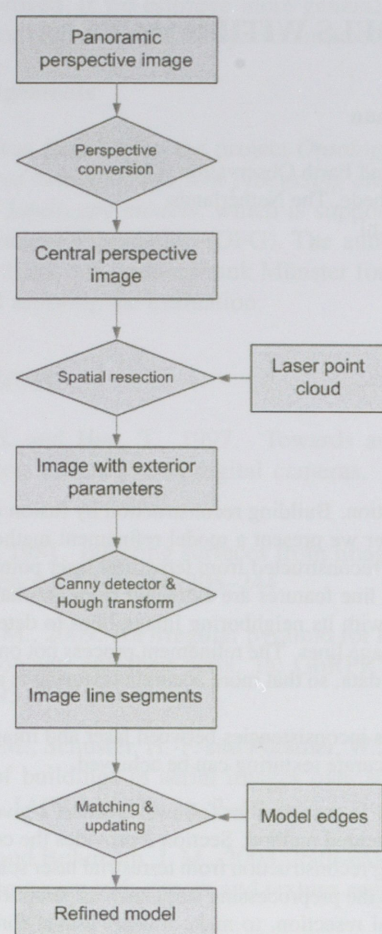


Figure 1: Model refinement process

3 BUILDING RECONSTRUCTION FROM TERRESTRIAL LASER SCANNING

Pu and Vosselman (2009) presents an automatic approach to extract building facade features from a terrestrial point cloud. The method first defines several important building features based on knowledge about building facades. Then the point cloud is segmented to planar segments. Finally, each segment is compared with building feature constraints to determine which feature this segment represents. The feature extraction method works fine for all facade features except for windows, because there are insufficient laser points reflected from window glass. Therefore a hole based window extraction method is introduced. Then polygons to extracted feature segments and the merging of polygons to a complete facade model. An advantage of this approach is that semantic feature types are extracted and linked to the resulting models, so that i) it is possible to get faster visualization by sharing the same texture for same feature type; ii) polygons can be associated with various attributes according to its feature type.

Figure 2 shows a building facade model which is reconstructed with the above approach. The generated building outline seems to coincide with laser points well. However, if we take a close look, it is easy to identify several mistakes from the model. By analyzing more models, we figured two main reasons for the modeling errors. They are:

- Limitations of outline generation method. For example, side wall's eave can "attract" the side boundary edges of the facade, and result in a slight wider polygon in horizontal di-

rection. The almost vertical or horizontal edges are forced to be vertical or horizontal; however, this is not always beneficial.

- Poor scanning quality. Due to the scanning strategy of static laser scanner, complete scanning of a scene seems impossible. There are always some parts which contain very sparse laser points, because of their visibility to any of the scan positions. Occluded zones without any laser points are also usual in laser point clouds. The lack of reference laser information leads to gaps in the final model. For example, the lower part of roofs are hardly scanned because the eaves occlude the laser beams. The directly fitted roof polygons are smaller than their actual sizes. Sometimes these gaps are foreseen and filled using knowledge. For example, we know a roof must attach to the upper side of an eave, so we can extend the roof polygon so that it intersects the eave. However, knowledge based estimation are not always correct.



Figure 2: A reconstructed building facade model, shown together with segmented laser points

4 PREPROCESSING

In order to extract straight lines, an image need to be in central perspective and undistorted. The exterior orientation parameters and focal length should be determined so that 3D model edges can be projected to the same image space for comparison. These are the two objectives of the preprocessing step. An omni-directional panoramic image called Cyclorama is used in our method development, therefore conversion of Cyclorama to central perspective are explained first in 4.1. A semi-automatic approach for exterior orientation calculation is given in 4.2.

4.1 Perspective conversion of Cyclorama

The Cycloramas are created from two fisheye images with a field of view of 185 degree each (van den Heuvel et al., 2007). The camera is turned 180 degree between the two shots. The Cycloramas we used contain image data for the full sphere stored in a panorama image of 4800 by 2400 pixels, corresponding to 360 degree in horizontal direction and 180 degree in vertical direction. Thus, on both directions the angular resolution is 0.075 degree per pixel. With the integrated GPS and IMU devices, all Cycloramas are provided with north direction aligned at $x=2400$ and horizontal plane aligned at $y=1200$.

The equiangular projection of the fisheye camera model is described in Schneider and Maas (2003). The projection of Cycloramas to central projective can be understood as projecting a

panoramic sphere part to an assumed plane. First, we make two lines by connect the image acquisition point (perspective center) with the most left and most right model vertices. The angles of the two lines with north direction derive the longitude boundaries of the region of interests (ROI). In practice we widen the ROI to both left and right by 100 pixels, because the image acquisition positions provided by GPS are not so reliable. The principal point is set on the sphere equator, with middle longitude of the two boundaries. Assuming the perspective center coincide in both perspectives, the pixels inside the ROI are converted from panoramic perspective to central perspective according to the following equations:

$$\alpha = \frac{x_p - x_0}{r} \quad (1)$$

$$\beta = \frac{y_p - y_0}{r} \quad (2)$$

$$\tan \alpha = \frac{x_c - x_0}{f} \quad (3)$$

$$\tan \beta = \frac{(y_c - y_0) \times \cos \alpha}{f} \quad (4)$$

where (x_p, y_p) is the pixel coordinate in panoramic perspective; (x_c, y_c) is the pixel coordinate in central perspective; (x_0, y_0) is the principal point; r is the angle resolution; α and β represent the longitude and latitude of the pixel on the panoramic sphere; f is the distance of the panoramic sphere center to the assumed plane, can also be seen as the focal length of the converted central perspective image. With equation 1 to 4 the unique relation between (x_p, y_p) and (x_c, y_c) can be determined.

4.2 Spatial resection

In order to get an unique solution for the six unknown exterior orientation parameters, at least observations of three image control points should be available to form 6 collinearity equations. Figure 3 illustrates the interface for selecting tie points from a laser point cloud and an image. In this implementation it is required to select at least four tie pairs, with one pair for error checking. If more than four pairs are selected, a least squares adjustment is performed to obtain better results.



edges should very well coincide with a strong image line. However, in practice there are often a small offset and an angle between a model edge and its corresponding image line. The optimal angle and distance threshold value are dependent on the quality of exterior orientation parameters and focal length.

2. A best match is chosen from all candidates according to either the collinearity of the candidates or the candidate's length (see the purple lines in Figure 5). It is a common case that a strong line is split to multiple parts by occlusions or shadows. If a number of Hough line segments belong to a same line, we set this line as the best match. If not, the longest candidate is chosen as the best match.



Figure 5: Matching model edges with image lines (Blue: model edges' projection in the image; red: Hough lines; green: candidates; purple: the best matches)

No spatial index is established in the image space to improve the comparison efficiency, because the search space is already localized to a single building facade, which includes only dozens of edges and Hough lines.

A limitation of this matching method is that it can hardly determine the correct corresponding edge if too many similar line features are within searching range. Simply comparing the geometry properties of position, direction and length are not sufficient in this case. For example, the eaves in Figure 5 result in many significant lines and they are all parallel and close to the wall's upper boundary edges. These eave lines can be distinguished if the eave is also reconstructed and included in the facade model, but ambiguity caused by pure color pattern is still difficult to be solved.

5.3 The refinement strategy

After matching, most model edges should be associated with a best matched image line. These model edges are updated by projecting to their best matched image line. There are some model edges which don't match any image lines. If no change is made to an edge with its previous or next edge changed, strange shapes like sharp corners and self-intersections may be generated. Therefore interpolations of the angle and distance change from the previous and next edges, are applied to the edges without matched image lines. With these refinement strategies, an original model

is updated to be consistent with the geometry extracted from images, and the model's geometry validity and general shape are also maintained.

Finally, the refined model edges in image space need to be transferred back to the model space. Because the model edges are only moved on their original 3D planes, which is known, the collinearity equations are used again to calculate the new 3D positions of all the modified model vertices.

6 TEST CASES

In this section, three data sets are experimented with the presented refinement method. The building models are produced with the reconstruction approach introduced in Section 3. All the images are originally provided as Cycloramas. The central perspective conversion and exterior orientation calculation follow the processes explained in Section 4.

6.1 The restaurant house

The inconsistencies between the model edges and image lines in Figure 6 are mainly due to inaccurate exterior orientation of the image. It is difficult to pick an image point accurately by manual operation. Picking the corresponding point in a laser point cloud is also a difficult job. Automated texturing of building facade models is desired in the context of our research. The quality of the exterior orientation is a key issue to the texturing effect. Even a minor inaccuracy in the exterior orientation parameters can lead to poor texture result, as shown in Figure 7(a). Applying our refinement method, several model edges are linked with their matched image lines (see Figure 6(b)), and are updated accordingly. The texture result is significantly improved as shown in Figure 7(b), with the sky's background color removed. However, the middle top part of the facade model is still not refined, because this image part is too blurred to output a Hough line.

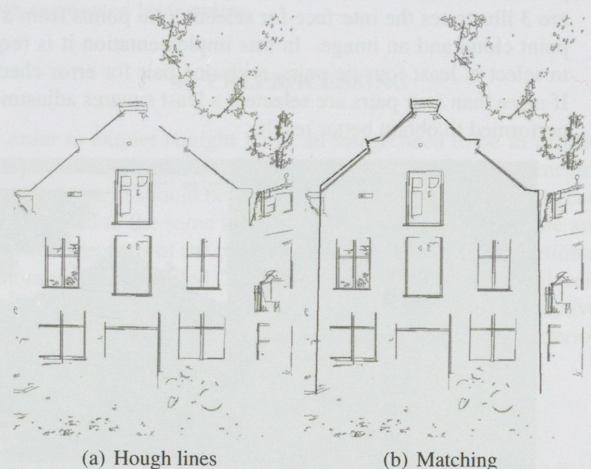


Figure 6: Matching model edges with image lines for refining the restaurant house's model

6.2 The town hall

The upper boundary of the town hall in Figure 8(a) contains a lot of tiny details, which are well recorded by laser scanning and modeled as sawtooth edges in the building facade model. Instead of adjusting the outline generation parameters in the reconstruction stage, we can also use the presented image based refinement to smooth the model outline. Figure 8(b) shows the matching step. The model's upper edges are successfully matched to the strong lines, which actually come from the eave. In this example

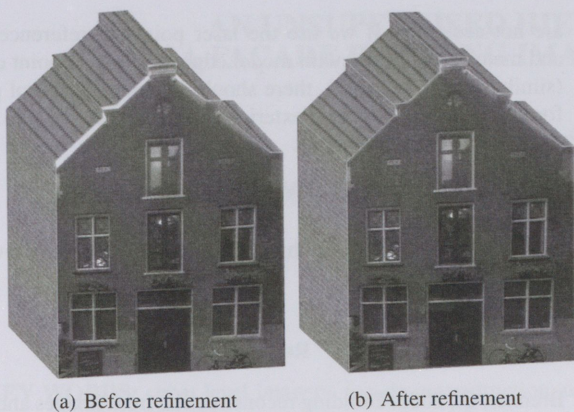


Figure 7: Comparison of textured restaurant house model before and after refinement

the wall's upper boundary is not detected in the image, because it is occluded from sunlight by the eaves when the image was taken.

The left boundary of the building model is modeled correctly from laser points by intersecting two large wall planes. However, in practice it is matched to a strong contrast caused by a water pipe on the wall (see Figure 8(b)). This again, reveals the limitations of this refinement method.

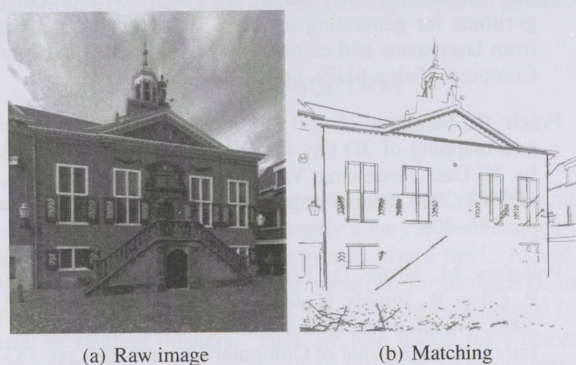


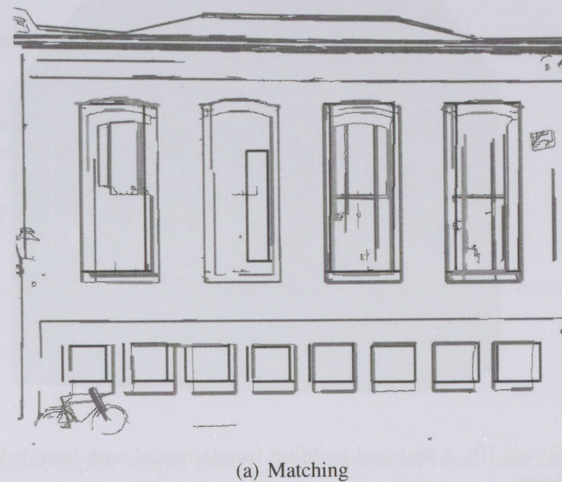
Figure 8: Matching model edges with image lines for refining the town hall's model

6.3 The wall with high windows

In this example, the refinement is applied to improve the windows extracted from the holes from laser points of a wall (Pu and Vosselman, 2009). The contrast of a window and its surrounding wall are usually rather obvious in optical data. Strong line features are frequently found at the windows' boundaries and frames, and can be used to refine the information from the laser altimetry. Figure 9(b) shows the window rectangles extracted from laser points, which contain a lot of errors due to limitation of segmentation and modeling algorithms. In Figure 9(a) we apply the same matching stretchy for refinement purpose, and the final result is shown in Figure 9(c). Most windows' boundaries are well corrected according to the image lines. The second left window in the upper row is not improved, because the difference between the modeled shape and the actual shape is too large to correlate them. There are some remaining errors, such as the first, third and sixth window (from left to right) in the lower row. This is because the parameters of Hough transform are too strict to generate any candidate line.

The textured final model after manual adjustment is shown in Figure 10. Without the refinement, 42 vertices need to be manually adjusted. Only 10 vertices need to be adjusted after the refinement. Note that the window rectangles are intruded inside the

wall plane, and the intrusion offset is derived from laser points reflected from window glass and window frames.



(b) Before refinement



(c) After refinement

Figure 9: Matching and refining window boundaries

6.4 Summary

The effectiveness as well as limitations of our refinement method are examined through the three test cases. We realize that the refining effect relies on the following prerequisites:

- Accurate exterior and interior orientations. In particular, the selected tie points for spatial resection should be sufficient (four or more), and should be distributed equally in both horizontal and vertical directions to minimize the computation error.
- No large occlusions in front of the building facade.

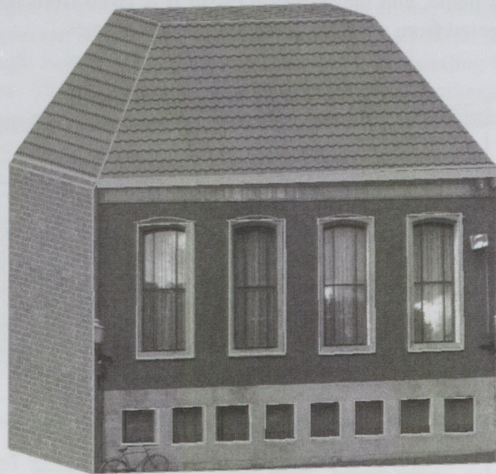


Figure 10: A textured building facade model with intruded windows

- Stronger contrast by geometries than optical factors (illumination, color pattern, etc.).

Some limitations of the current refinement method have been located at:

- It cannot solve ambiguities caused by multiple lines with similar geometry properties.
- It cannot distinguish whether a model-image inconsistency is caused by reconstruction errors or inaccurate exterior orientation.

Knowledge based reasoning of the image information is the key to the first problem. The current matching stretchy is rather local. Experiments show that the offset direction between the model edges and their matched image lines are mostly same, which is obviously caused by inaccurate exterior orientation. A globe matching process (RANSAC over offsets for example) should be able to estimate the correct exterior orientations.

7 CONCLUSIONS AND OUTLOOK

In this paper we present a model refinement method, which uses the lines extracted from close-range images to improve building models reconstructed from terrestrial laser point clouds. With the refinement, several modeling errors caused by either gaps in laser data or reconstruction algorithm, are corrected with image information. Texturing is also improved after the refinement.

Nowadays it is more and more common for acquisition platforms to acquire laser data and optical data simultaneously. Line extraction from images is very accurate, while laser points are more suitable to extract planar features. Efficient fusing of laser points and image naturally avoids many barriers for building reconstruction from either sides. The attempt through our refinement method shows promising future for automated building reconstruction by fusing laser altimetry and optical methods.

Two directions of the future work: knowledge based image reasoning and global matching, have been suggested earlier. Besides, nowadays the mainstream image acquisition systems usually determine exterior orientations via GPS and IMU, but they

are not accurate. If we use the laser points as reference data, and match image lines with model edges from laser point clouds (similar to this research), there should be enough control points for estimating the accurate exterior orientations for images.

ACKNOWLEDGEMENTS

The authors would like to thank Cyclomedia B.V. for providing the Cyclorama data.

References

- Brenner, C., 2005. Building reconstruction from images and laser scanning. *International Journal of Applied Earth Observations and Geoinformation* 6(3-4), pp. 187-198.
- Canny, J., 1986. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence* pp. 679-698.
- Dick, A., Torr, P., Ruffle, S. and Cipolla, R., 2001. Combining single view recognition and multiple view stereo for architectural scenes. In: *Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings, Vol. 1*.
- Frueh, C., Jain, S. and Zakhori, A., 2005. Data processing algorithms for generating textured 3D building facade meshes from laser scans and camera images. *International Journal of Computer Vision* 61(2), pp. 159-184.
- Frueh, C., Sammon, R. and Zakhori, A., 2004. Automated texture mapping of 3D city models with oblique aerial imagery. In: *3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium on*, pp. 396-403.
- Pollefeys, M., Nister, D., Frahm, J., Akbarzadeh, A., Mordohai, P., Clipp, B., Engels, C., Gallup, D., Kim, S., Merrell, P. et al., 2008. Detailed real-time urban 3d reconstruction from video. *International Journal of Computer Vision* 78(2), pp. 143-167.
- Pu, S. and Vosselman, G., 2009. Knowledge based reconstruction of building models from terrestrial laser scanning data (in press). *ISPRS J. Photogramm. Remote Sens.*
- Schindler, K. and Bauer, J., 2003. A model-based method for building reconstruction. In: *First IEEE International Workshop on Higher-Level Knowledge in 3D Modeling and Motion Analysis, 2003. HLK 2003*, pp. 74-82.
- Schneider, D. and Maas, H., 2003. Geometric modelling and calibration of a high resolution panoramic camera. *Optical 3D Measurement Techniques VI* (Eds. Gruen, A. and Kahmen, H.) 2, pp. 122-129.
- van den Heuvel, F., Verwaal, R. and Beers, B., 2007. Automated Calibration of Fisheye Camera Systems and the Reduction of Chromatic Aberration. *Photogramm. Fernerkund. Geoinf.* 3, pp. 157.
- Vosselman, G., 2002. Fusion of laser scanning data, maps, and aerial photographs for building reconstruction. In: *2002 IEEE International Geoscience and Remote Sensing Symposium, 2002. IGARSS'02, Vol. 1*.