REFINEMENT OF BUILDING FASSADES BY INTEGRATED PROCESSING OF LIDAR AND IMAGE DATA

Susanne Becker, Norbert Haala

Institute for Photogrammetry (ifp), University of Stuttgart, Germany
Geschwister-Scholl-Straße 24D, D-70174 Stuttgart
Forename.Lastname@ifp.uni-stuttgart.de

KEY WORDS: Three-dimensional, Point Cloud, Urban, LIDAR, Modelling, Façade Interpretation

ABSTRACT:

Urban models extracted from airborne data have to be refined for tasks like the generation of realistic visualisations from pedestrian viewpoints. Within the paper, terrestrial LIDAR data as well as façade imagery is used to increase the quality and amount of detail for the respective 3D building models. These models as they are available from airborne data collection provide a priori information, which can be integrated efficiently both for the georeferencing of the terrestrial data and the subsequent geometric refinement. After alignment of the terrestrial data to the given 3D model, window structures are first extracted approximately from the LIDAR point clouds. These structures are then further refined by 3D edges which are extracted from the overlapping façade images. Our modelling process applies a 3D object representation by cell decomposition, which can be used efficiently for building reconstruction at different scales.

1. INTRODUCTION

The area covering collection of urban models is usually based on the evaluation of aerial data like stereo images or LIDAR. The available algorithms provide 3D building representations which are sufficient for applications like simulations and visualisations at small or medium scale. However, for large scale applications like the generation of very realistic visualisations from pedestrian viewpoints, the quality and amount of detail for urban models from aerial data has to be improved. As an example, due to the viewpoint restrictions of airborne platforms, detailed information for the facades of the buildings frequently is not available. Thus, to improve the visual appearance of the buildings, terrestrial images are often mapped against the facades. However, this substitution of geometric modelling by real world imagery is only feasible to a certain degree. For instance, protrusions at balconies and ledges, or indentations at windows will disturb the visual impression of oblique views. Thus, geometric refinement is still necessary for a number of applications.

In order to enable the geometric modelling of building facades, either terrestrial laser scanning or image measurement can be used. Within this paper, the potential of these data sets for facade interpretation is demonstrated exemplarily for the extraction of window objects. In our opinion, an image based approach like it is for example presented by (Mayer & Reznik, 2006) considerably profits from the additional availability of densely sampled point clouds from terrestrial laser scanning. An integrated collection of such data sets is feasible by mobile systems, where a laser scanner and a camera are mounted on a car. Such a system was for example applied by (Früh & Zakhor, 2003) to generate textured meshes for visual representation of building facades. In our investigations, standard equipment consisting of a digital camera and a terrestrial laser scanner is used. To avoid data collection from scratch and to facilitate both the georeferencing and the modelling process, existing

building models as they are provided from airborne data collection are closely integrated to all steps. Thus, we aim at a data driven geometric enrichment of building facades, whereas approaches using grammar based façade descriptions are more likely to focus on semantic modelling and interpretation (Brenner & Ripperda, 2006), (Alegre & Dallaert, 2004).

In contrast to other approaches based on building representations by constructive solid geometry (CSG) or boundary representation (B-Rep), we apply a representation of the buildings by cell decomposition. By these means, the problems to correctly generate topologically correct boundary representations can be avoided. The same holds true if geometric constraints such as meeting surfaces, parallelism and rectangularity have to be met. The formulation of such regularization conditions is also simplified if an object representation based on CSG is used. However, while CSG is widely used in computer aided design since it allows for powerful and intuitive modelling (Mäntylä, 1988), most visualization and simulation applications require the additional derivation of a boundary representation. While this is conceptually easy, its correct and efficient implementation can be difficult. Problems can arise from error-prone measurements, limited numerical precision and unstable calculation of intersections.

These problems are facilitated by the concept of cell decomposition. Similar to CSG, complex solids are described by a combination of relatively simple, basic objects in a bottom up fashion. In contrast to CSG, which combines simple primitives by means of regularized Boolean set operators, decomposition models are limited to adjoining primitives. Since the basic primitives must not intersect, they are thus 'glued' together to get the final model. In this sense, cell decomposition is similar to a spatial occupancy enumeration, where the object space is subdivided by non overlapping cubes of uniform size and orientation. Nevertheless, cell decompositions are based on a variety of basic cells, which may be any objects that are

topologically equivalent to a sphere i.e. do not contain holes. This allows for a simplified combination of the respective elements, while the disadvantages of exhaustive enumeration like large memory consumption and the restricted accuracy of the object representation can be avoided.

Since it is a prerequisite for further processing, the georeferencing process of the collected images and LIDAR data is described in Section 2. For this purpose, the collected data sets are aligned to the existing buildings from airborne data collection. The geometric refinement of the facades presented in Section 3 is implemented as a two-step approach. In order to integrate window objects to the existing coarse building model, cell decomposition is used. First, the windows and doors are modelled from the terrestrial LIDAR data, while the window frames are further refined by photogrammetric analysis of the images in a subsequent step.

2. DATA PREPARATION AND ALIGNMENT

The combined evaluation of the terrestrial LIDAR and image data requires the co-registration of the different data sets as a first processing step. The alignment of single images to a triangulated 3D point cloud can for example be realised based on corresponding linear primitives provided by a suitable edge detection process (Haala & Alshawabkeh, 2006). In our application, approximate geometry of the respective buildings is already available and can therefore be used to facilitate the georeferencing process. The quality and amount of detail of this data set is typical for such 3D models, like they are available area covering for a number of cities. Our exemplary 3D city model, which is maintained by the City Surveying Office of Stuttgart, features roof faces collected semi-automatically by photogrammetric stereo measurement. In contrast, the outlines of the buildings were captured by terrestrial surveying. Thus, the horizontal position accuracy of façade segments, which were generated by extrusion of this ground plan, is relatively high, despite the fact that they are limited to planar polygons.

2.1 Georeferencing of LIDAR data

During the collection of the 3D point clouds, a low-cost GPS and a digital compass were mounted on top of the used HDS 3000 laser scanner to allow for a direct georeferencing of the terrestrial scans. This approximate solution is then refined by an automatic registration of the laser scans against the 3D building model using a standard iterative closest point (ICP) algorithm (Böhm & Haala, 2005).
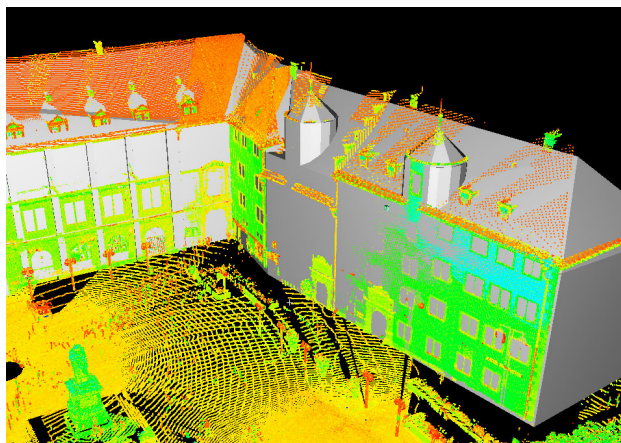

Figure 1: 3D point cloud from laser scanning aligned with a virtual city model.

As it is demonstrated in Figure 1, after this step the 3D point cloud and the 3D city model are available in a common reference system. Thus, relevant 3D point measurements can be selected for each building façade by a simple buffer operation. These 3D points are then transformed to a local coordinate system as defined by the façade plane. Figure 2 shows the resulting point cloud, which has an approximate spacing of 4cm.
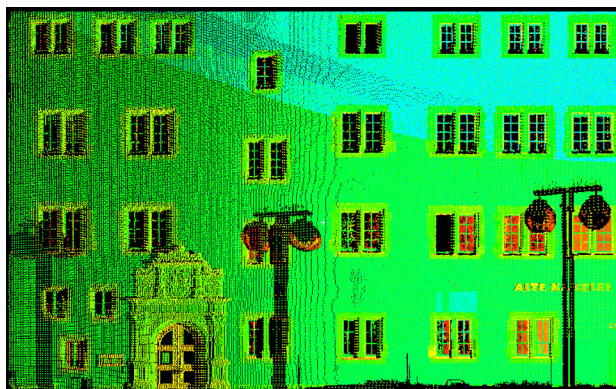

Figure 2. 3D point cloud as used for the geometric refinement of the corresponding building façade.

Since the LIDAR measurements are more accurate than the available 3D building model, the final reference plane is determined from the 3D points by a robust estimation process. After mapping of the 3D points to this reference plane, further processing can be simplified to a 2.5D problem. As an example, while assuming that the refined geometry of the façade can be described sufficiently by a relief, the differences between the measured 3D laser points and the given façade polygon can be interpolated to a regular grid.

2.2 Alignment of image data

Image orientation is the first step within the photogrammetric 3D modelling. Usually bundle adjustment is the method of choice if accurate orientation parameters are to be estimated. The determination of initial orientation parameters by spatial resection requires control points that can be obtained from the images and the LIDAR point cloud. Additionally, tie points are necessary for connecting the images. In the recent past, much effort has been made to develop approaches that automatically extract such tie points from images of different types (short, long, and wide baseline images) (Remondino & Ressl, 2006). While matching procedures based on cross-correlation are well suited for short baseline configurations, images with a more significant baseline are typically matched by means of interest points. However, these techniques would fail in case of wide baseline images acquired from considerably different viewpoints. This is due to big perspective effects that are caused by the large camera displacement. Points and corners cannot be reliably matched. Thus, interest point operators have to be replaced by region detectors and descriptors. As an example, the Lowe operator (Lowe, 2004) has been proved to be a robust algorithm for wide baseline matching (Mikolajczyk & Schmid, 2003).

Figure 3 shows the image data from a calibrated camera (NIKON D2x Lens NIKKOR 20mm). For the automatic provision of tie points the SIFT (scale invariant feature transform) operator has been applied to extract and match keypoints. Wrong matches are removed by a RANSAC based estimation (Fischler & Bolles, 1981) of the epipolar geometry using Nister's five point algorithm (Nister, 2004).

Figure 3. Image data for photogrammetric modelling.

The control points for the final bundle adjustment, which is performed with the Australis software package, are measured manually in the images and the 3D laser points.

3. FAÇADE RECONSTRUCTION

The reconstruction algorithm presented in this paper is a two-step approach based on terrestrial LIDAR and image data. It aims at the geometric facade refinement of an existing coarse building model by the integration of window objects. At first, cell decomposition is used to model windows and doors from the LIDAR data. In a second step, the window frames are further refined by photogrammetric analysis of the images.

3.1 Façade Refinement By Terrestrial LIDAR

The idea of the first part of our reconstruction algorithm is to segment a 3D object with a flat front face into 3D cells. Each 3D cell represents either a homogeneous part of the façade or a window area. Therefore, they have to be differentiated based on the availability of measured LIDAR points. After this classification step, window cells are eliminated while the remaining façade cells are glued together to generate the refined 3D building model. The difficulty is finding planar delimiters from the LIDAR points that generate a good working set of cells. Since our focus is on the reconstruction of the windows, the planar delimiters have to be derived from the 3D points that were measured at the window borders. These points are identified by a segmentation process.

3.1.1 Cell Generation

Point cloud segmentation. As it is visible for the façade in Figure 2, usually fewer 3D points are measured on the façade at window areas. This is due to specular reflections of the LIDAR pulses on the glass or points that refer to the inner part of the building and were therefore cut off in the pre-processing stage. If only the points are considered that lie on or in front of the façade, the windows will describe areas with no point measurements. Thus, our point cloud segmentation algorithm detects window edges by these no data areas. In principle, such holes can also result from occlusions. However, this is avoided by using point clouds from different viewpoints. In that case, occluding objects only reduce the number of LIDAR points since a number of measurements are still available from the other viewpoints.

During the segmentation process, four different types of window borders are distinguished: horizontal structures at the top and the bottom of the window, and two vertical structures that define the left and the right side. For instance, the edge points of a left window border are detected if no neighbour measurements to their right side can be found in a pre-defined search radius at the façade plane. The search radius should be set to a value a little higher than the scan point distance on the façade. The extracted edge points are shown in Figure 4. While most of the edge points can be correctly identified this way, the algorithm often fails to find points at window corners.

However, this is not a real problem, as long as there are enough points to determine the window borders. For this purpose, horizontal and vertical lines are estimated from non-isolated edge points. The resulting set of window lines is depicted in Figure 5.
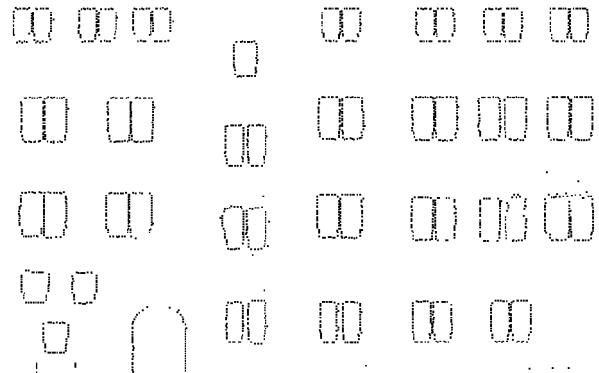

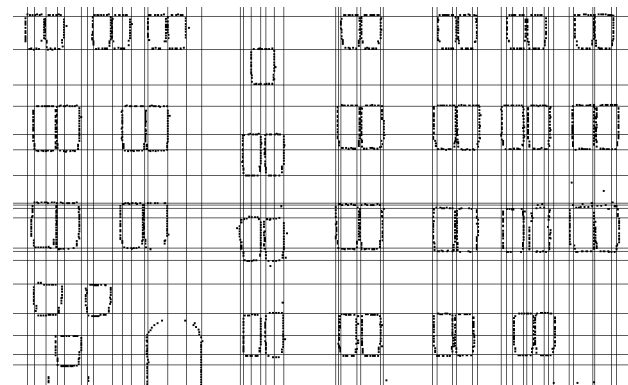Figure 4. Detected edge points at horizontal and vertical window structures.

Figure 5. Detected horizontal and vertical window lines.

Spatial-Partitioning. Each boundary line defines a partition plane, which is perpendicular to the building façade. For the determination of the window depth, an additional partition plane is estimated from the LIDAR points measured at the window crossbars. These points are detected by searching a plane parallel to the façade, which is shifted in its normal direction. The set of all partition planes provides the structural information for the cell decomposition process. Therefore, it is used to intersect the existing building model producing a set of small 3D cells.

3.1.2 Classification of 3D cells

In a next step, the generated 3D cells have to be classified into building and non-building fragments. For this purpose, a 'point-availability-map' is generated. It is a binary image with low resolution where each pixel defines a grid element on the façade. The optimal size of the grid elements is a value a little higher than the point sampling distance on the facade.

As it can be seen in Figure 6, black pixels are raster elements where LIDAR points are available, while white pixels represent grid elements with no 3D point measurements. Of course, the already extracted edge points in Figure 4 and the resulting structures in Figure 5 are more accurate than the rasterized point-availability-map. However, this limited accuracy is acceptable since the binary image is only used to classify the 3D cells, which are already created from the detected horizontal and vertical window lines. This is implemented by computing

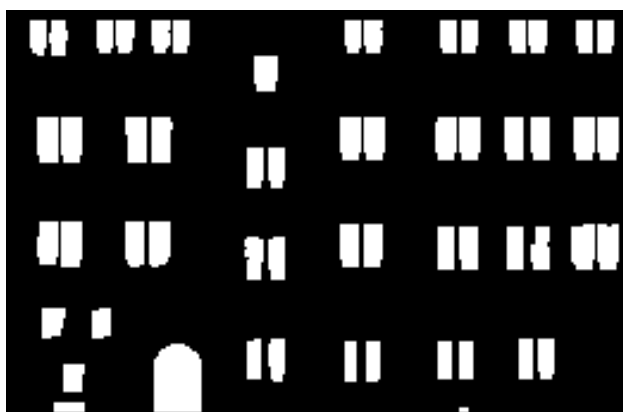the ratio of façade to non-façade pixels for each generated 3D cell.

Figure 6. Point-availability-map.

As a consequence of the relative coarse rasterization and the limited accuracy of the edge detection, the 3D cells usually do not contain facade pixels or window pixels, exclusively. Within the classification, 3D cells including more than 70% façade pixels are defined as façade solids, whereas 3D cells with less than 10% façade pixels are assumed to be window cells. These segments are depicted in Figure 7 as grey (façade) and white (window) cells.
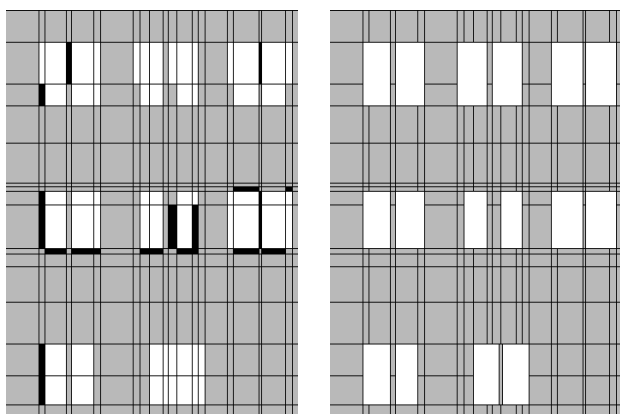

Figure 7. Classification of 3D cells before (left) and after enhancement (right).

Classification Enhancements. While most of the 3D cells can be classified reliably, the result is uncertain especially at window borders or in areas with little point coverage. Such cells with a relative coverage between 10% and 70% are represented by the black segments in the left of Figure 7. For the final classification of these cells, neighbourhood relationships as well as constraints concerning the simplicity of the resulting window objects are used. As an example, elements between two window cells are assumed to belong to the façade, so two small windows are reconstructed instead of one large window. This is justified by the fact that façade points have actually been measured in this area. Additionally, the alignment as well as the size of proximate windows is ensured. For this purpose, uncertain cells are classified depending on their neighbours in horizontal and vertical direction. Within this process, it is also guaranteed that the merge of window cells will result in convex window objects. Figure 7 (right) illustrates the enhanced classification result.
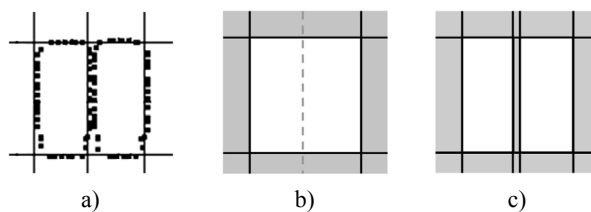

a)   b)   c)

Figure 8. Integration of additional façade cell.

As it is depicted in Figure 8, additional façade cells can be integrated easily if necessary. Figure 8a shows the LIDAR measurements for two closely neighboured windows. Since in this situation only one vertical line was detected, a single window is reconstructed (Figure 8b). To overcome this problem, the window object is separated into two smaller cells by an additional façade cell. This configuration is kept if façade points are available at this position (Figure 8c).

3.1.3 Façade Modelling

Within the following modelling process, the window cells are cut out from the existing coarse building model. The result of the building façade reconstruction is given in Figure 9. The front of the pyramidal wall dormer is not considered as being a part of the façade. Therefore, the reconstruction approach is applied on the roof extension, separately.
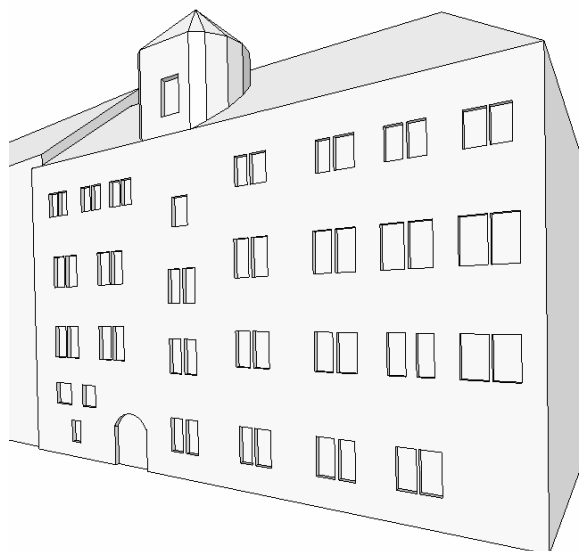

Figure 9. Refined facade of the reconstructed building.

While the windows are represented by polyhedral cells, also curved primitives can be integrated in the reconstruction process. This is demonstrated exemplarily by the round-headed door of the building. Furthermore, our approach is not limited to the modelling of indentations like windows or doors. Details can also be added as protrusions to the façade. LIDAR points that are measured at protrusions can be detected easily since they are not part of the façade plane but lying in front of it. If these points are classified as non-façade points, protrusion areas can be identified in the same way as window regions, just by searching no data areas within the set of points that belong to the facade. The availability of LIDAR points in front of the façade helps to classify the derived 3D cells as protrusion cells. Their extent in the façade's normal direction can be reconstructed by fitting planes to the measured protrusion points.

3.2 Facade Refinement By Photos

The level of detail for 3D objects that are derived from terrestrial laser scanning is limited depending on the point sampling distance. Small structures are either difficult to detect or even not represented in the data. By integrating image data in the reconstruction process the amount of detail can be increased. This is exemplarily shown for the reconstruction of window crossbars.

3.2.1   Derivation of 3D edges

Having oriented the image data, 3D information can be derived from corresponding image features in order to reconstruct details of the façade such as crossbars. For this purpose, edge points are extracted from the images by a Sobel filter. These edge point candidates are thinned and split into straight segments. Afterwards, the resulting 2D edges of both images have to be matched. However, frequently occurring façade structures, such as windows and crossbars, hinder the search for corresponding edges. Therefore, the boundaries of the windows that have already been reconstructed from the LIDAR points are projected into both images. Only the 2D edges lying inside these image regions are considered for the following matching process. Thus, possible mismatches are reduced, even though, they cannot be avoided entirely. Figure 10 depicts the selected 2D edges for an exemplary window in both images.

Figure 10. Selected 2D edges for a window in both images.

Remaining false correspondences lead to 3D edges lying outside the reconstructed window. Therefore, these wrong edges can be easily identified and removed. In addition, only horizontal and vertical 3D edges are considered for the further reconstruction process. The reconstructed wrong (green) and correct (red) 3D edges are shown in local façade coordinates in Figure 11. The position of the window that has been derived from the LIDAR data is illustrated in black.
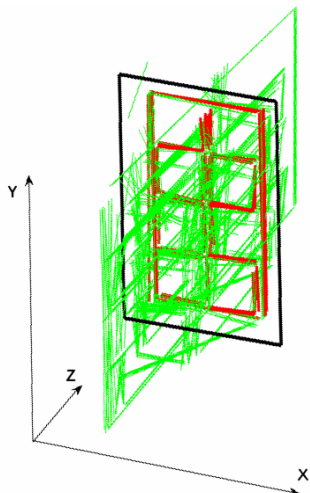

Figure 11. Wrong (green) and correct (red) 3D window edges.

3.2.2   Reconstruction of additional façade structures

Photogrammetric modelling allows the extraction of well-defined image features like edges and points with high accuracy. By contrast, points from terrestrial laser scanning are measured in a pre-defined sampling pattern, unaware of the scene to capture. That means that the laser scanner does not explicitly capture edge lines, but rather measures points at constant intervals. Furthermore, laser measurements at edges and corners may provide erroneous and unpredictable results because of the laser beam split that is caused at the object border. For these reasons, the positional accuracy of window borders that are reconstructed from LIDAR points is limited compared to the photogrammetrically derived 3D edges at crossbars. As a consequence, the 3D reconstructions from laser points and images may be slightly shifted. Therefore, the reconstruction of the crossbars is done as follows:

For each window, hypotheses about the configuration of the crossbars are generated and tested against the 3D edges derived from the images. Possible shapes are dynamically generated as templates by recursively dividing the window area in two or three parts. Recursion stops when the produced glass panes are too small for a realistic generation of windows. The minimum width and height of the glass panes are restricted by the same threshold value. After each recursion step, the fitting of the template with the 3D edges is evaluated. The partition is accepted if 3D edges are available within a buffer area around the dividing line. In a final step, the crossbars and the window frame are modelled. For this purpose, new 3D cells with a pre-defined thickness are generated at the accepted horizontal and vertical division lines as well as at the window borders. The result is exemplarily shown for two windows in Figure 12.
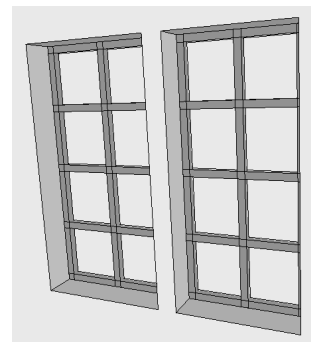

Figure 12. Reconstructed crossbars for two windows.

While most of the crossbars can be reconstructed reliably, problems may arise for windows that are captured under oblique views. Perspective distortions or occlusions make it difficult to detect 2D edges at crossbars (Figure 13). Consequently, fewer 3D edges can be generated thereof in those areas.
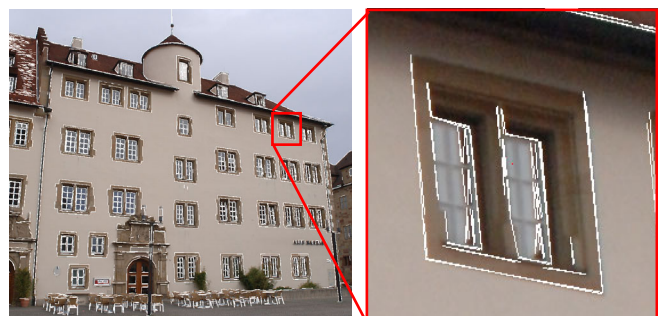

Figure 13. Detected 2D edges for a window captured under an oblique view.

To overcome this problem, neighbourhood relationships are taken into account within the final modelling step. The crossbar configuration is assumed to be equal for all windows of similar size which are located in the same row or column. Based on this assumption, similar windows can be simultaneously processed. Thus, the crossbar reconstruction leads to robust results even for windows that are partially distorted or feature strong perspective distortions in the respective image areas.
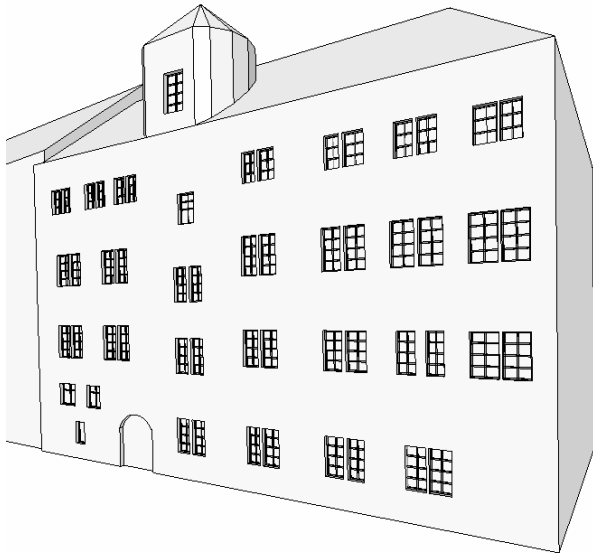

Figure 14. Refined facade with detailed window structures.

Figure 14 shows the final result of the building façade reconstruction from terrestrial LIDAR and photogrammetric modelling. This example demonstrates the successful detection of crossbars for windows of medium size. However, the dynamic generation of templates even allows for the modelling of large window areas as they often occur at facades of big office buildings.

4. CONCLUSION

Within the paper, an approach for the refinement of 3D building models based on cell decomposition was presented. As it was already proved for the automatic generation of topologically correct building models at different levels of detail (Haala et al, 2006), this approach allows the simple integration and removal of geometric detail for given building models. Even more important, symmetry relations like coplanarity or alignment can be guaranteed even for larger distances between the respective building parts. Thus, despite of the limited extent of the window primitives, which were extracted by the analysis of terrestrial LIDAR and images, structural information can be generated for the complete building. In principle, this information can then be used to support the façade interpretation at areas where measurements are only available with reduced quality and reliability. For these reasons, in our opinion this approach has a great potential for processes aiming at the reconstruction and refinement of building models from multiple data sets.

5. REFERENCES

Alegre, F. & Dallaert, F., 2004. A Probabilistic Approach to the Semantic Interpretation of Building Facades. International Workshop on Vision Techniques Applied to the Rehabilitation of City Centres, pp. 1-12.

Böhm, J. & Haala, N., 2005. Efficient Integration of Aerial and Terrestrial Laser Data for Virtual City Modeling Using LASERMAPS. IAPRS Vol. 36 Part 3/W19 ISPRS Workshop Laser scanning, pp.192-197.

Brenner, C. & Ripperda, N., 2006. Extraction of Facades Using RjMCMC and Constraint Equations. *Remote Sensing and Spatial Information Sciences* (36) 3.

Fischler, M. A. & Bolles, R. C., 1981. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, Vol. 24, pp. 381-395.

Früh, C. & Zakhor, A., 2003. Constructing 3D City Models by Merging Ground-Based and Airborne Views. *IEEE Computer Graphics and Applications,* Special Issue Nov/Dec.

Haala, N. & Alshawabkeh, Y., 2006. Combining Laser Scanning and Photogrammetry - A Hybrid Approach for Heritage Documentation. The 7th International Symposium on Virtual Reality, Archeology and Cultural Heritage VAST, pp.163-170.

Haala, N., Becker, S. & Kada, M., 2006. Cell Decomposition for the Generation of Building Models at Multiple Scales. IAPRS Vol. XXXVI Part III, Symposium Photogrammetric Computer Vision, pp. 19-24.

Lowe, D., 2004. Distinctive image features from scale-invariant keypoints. IJCV, Vol. 60(2), pp. 91-110.

Mäntylä, M., 1988. *An Introduction to Solid Modeling*. Computer Science Press, Maryland, U.S.A.

Mayer, H. & Reznik, S., 2006. MCMC Linked With Implicit Shape Models and Plane Sweeping for 3D Building Facade Interpretation in Image Sequences. IAPRS Vol. XXXVI, Part. 3.

Mikolajczyk, K. & Schmid, C., 2003. A performance evaluation of local descriptors. Proc. Conf. Computer Vision and Pattern Recognition, pp. 257-264.

Nistér, D., 2004. An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 26(6), pp. 756-770.

Remondino, F. & Ressl, C., 2006. Overview and experiences in automated markerless image orientation. IAPRSSIS, Vol. 36, Part 3, pp.248-254.