

WINDOW DETECTION IN SPARSE POINT CLOUDS USING INDOOR POINTS

S. Tuttas, U. Stilla

Photogrammetry and Remote Sensing, Technische Universität München, 80290 München, Germany -
sebastian.tuttas@bv.tum.de, stilla@tum.de

KEY WORDS: point cloud, building, city, façade reconstruction, airborne laser scanning

ABSTRACT:

This paper describes an approach for detecting windows from multi-aspect airborne laser scanning point clouds which were recorded in a forward looking view. Since the resolution of the point cloud is much lower than from terrestrial laser scanning, new methods have to be developed to detect and, in a further step, reconstruct façade structures. The façade planes are detected using point normals and a regiongrowing algorithm. The approach for window detection uses the points which are lying behind the detected façades planes (indoor points). Regularities in the appearance of these points are of special interest to enable the detection of windows which are only weakly represented in the point cloud. Therefore it is checked with a Fourier Transform if a repetitive structure can be extracted. Otherwise peaks in the density of the indoor points are used to detect the windows. The approach is tested on data from four overflights over the area around the TU München. The tests show that windows having a repetitive structure can be detected well for larger façade parts which provide enough samples but the approach shows deficits for small façade parts and in the case of disturbing intrusions.

1. INTRODUCTION

1.1 Motivation

3D-city models are used for several applications, for example urban planning, navigation or visualisation of buildings of touristic interest. In these cases it is sufficient to use polyhedral models. They can be used with or without texture. But there are also applications which require a more detailed façade reconstruction. To analyse Persistent Scatterers for radar image interpretation façade details are of special interest (Auer et al. 2011), for energetic assessment of buildings with thermal cameras the area of the façade without the window regions is needed (Iwaszczuk et al. 2011). Also the visualisation of a building can be made more realistic by modelling geometric structures like windows and doors.

In most cases the polyhedral models are received from airborne laser scanning data. From this the roof structure can be modelled but only less points can be found on the façade.

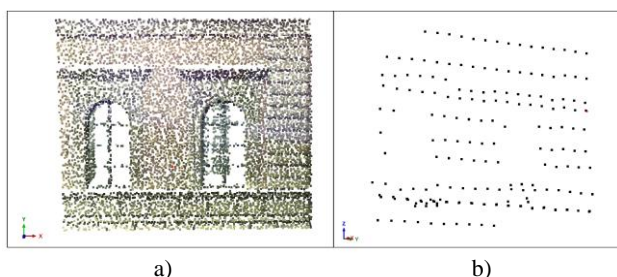


Figure 1. Example for laser point clouds: a) Dense point cloud (~ 200 points / m^2) from terrestrial laser scanning. b) sparse point cloud (~ 5 points / m^2) from forward looking airborne laser scanning

Façade reconstruction methods make usually use of terrestrial laser scanning data, e.g. from a street mapper. This data has often a very high point density, but lacks of roof data. Normally

only the part of the façade can be seen, which is oriented towards the street.

Airborne multi-aspect laser scanning data makes it possible to reconstruct building façades and roofs for entire buildings from a single data set. This data is a compromise between completeness and point density, which cannot be as high as from terrestrial scanners. In Figure 1 the point density of a terrestrial point cloud and from our test data is compared.

1.2 Related Work

A comprehensive overview on 3D building reconstruction from LiDAR and from image data is given by Haala and Kada (2010). Their paper has two main parts, one describes approaches for roof shape reconstruction, the other one outlines approaches on building façades. They state a large variety of different works on the topic of roof shape reconstruction, whereby they distinguish between three main groups: reconstruction with parametric shapes, reconstruction based on segmentation and reconstruction by DSM simplification.

Since the roof reconstruction is more advanced and the different approaches are well described in Haala and Kada (2010) only papers concentrating on façade reconstruction are presented in the following.

A basic problem is always the detection of structures in the point cloud, mainly planes. An overview on this topic is given by Vosselmann et al. (2004).

Ripperda (2008) derived grammar rules for façade parameters from images which can be used in the reconstruction process using a Reversible Jump Markov Chain Monte Carlo method. Also Becker (2009) uses a formal grammar to reconstruct building façades. The grammar is derived from terrestrial laser scanning data and is refined with image data. With the help of the grammar also building parts, which are occluded, can be modelled.

Boulaassal et al. (2009) detect window contours using a 2D-Triangulation of the façade plane. Since the windows are represented by holes in the façade, it is searched for the longest triangle sides to find the points surrounding the windows.

Schmittwilken and Plümer (2010) use a model-based reconstruction approach. They use training data to create probability density functions for the shape parameters of windows, doors and stairs which are used for a prefiltering of the point cloud. The selection of the most likely sample for a certain object structure is done by an adapted RANSAC approach which uses a more efficient criterion for the scoring. Pu and Vosselman (2009) extract features from a segmented point cloud by defining constraints for the different façade features. They also use a hole-based window extraction method using a TIN. Knowledge is brought in to complete the parts of the building, which are occluded.

1.3 Concept

The approaches mentioned before have all in common that they use high resolution point clouds (hundred to several hundred points per square meter). Our data has only approximately 5 points per square meter, but there can also be parts with a total lack of points. Because of the oblique view the point density on the façade can vary for a single scan. This makes it hard to use holes in the point data for window detection, what lead us to an approach which uses regular patterns of points behind the façade (= indoor points). The proposed workflow can be found in Figure 2.

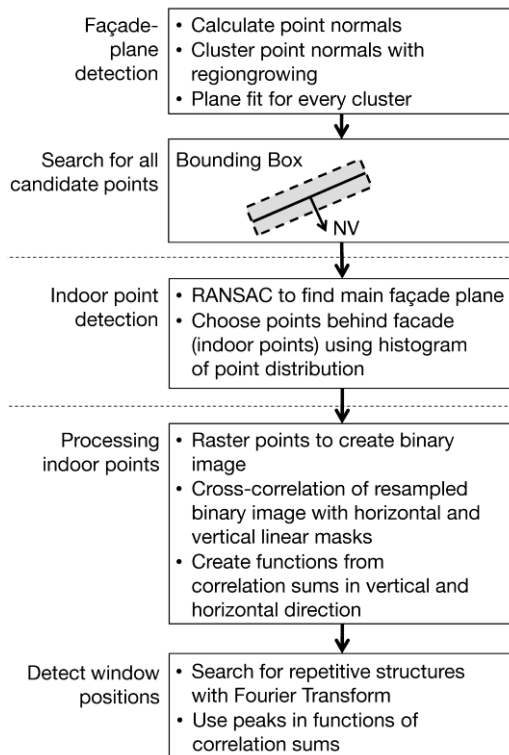


Figure 2. Flowchart of the presented approach

The approach is based on two basic assumptions:

- Laser pulses can pass through windows and are reflected inside the building. These points are a few cm to a few m behind the main façade plane.
- Windows are often arranged in a regular way, at least in one façade direction (vertical or horizontal). This also means that it is likely that a window exists at the same position on each floor.

First façade planes have to be detected, what is described in Section 2. After that every façade is processed on its own to detect the windows, what is delineated in Section 3. This has three parts. The first one is the detection of the indoor points (3.1), what is done by fitting a Gaussian function to a histogram of the point distribution. Then the indoor points are rastered to generate a binary image. This is cross-correlated with a horizontal and a vertical line (3.2). Finally it is searched for repetitive structures in the resulting correlation images (3.3). The approach is tested on a data set of the TU München, what is shown in Section 4. Conclusions and outlook are given in the last section.

2. FAÇADE PLANE DETECTION

First normals for every point have to be calculated using a search radius r depending on the point cloud density (e.g. 3 m). The normals are used to find potential façade points. All points having normals which deviate more than $\Delta\varphi$ from the horizontal plane are rejected. This helps to reduce the points, which have to be processed. The point normals can be used in different ways to support the segmentation process. For example Awwad et al. (2010) improved a RANSAC algorithm by including a check between the normal vector of the point cloud and the hypothesised RANSAC plane. Here a regiongrowing algorithm using the normals is applied to extract the façade planes from the point cloud (see Figure 3). For every point always the n nearest neighbours are considered. The points are allocated to the same segment if the distance between the points is less than Δd (distance threshold) and if the angle between the normal vectors, projected into the horizontal plane, is less than $\Delta\delta$ (angle threshold).

```

1: segk = 1
2: for i = 1 : # points
3:   get n Nearest Neighbours (NN)
4:   calculate distance  $d$  and angle  $\delta$  between normal vectors
5:   if  $d < \Delta d$  &&  $\delta < \Delta\delta$ 
6:     find best NN (NNbest)
7:     if seg(NNbest) = ∅ || seg(seg(Pi)) = ∅
8:       if seg(Pi) = x: seg(NNbest) = x
9:       elseif: seg(NNbest) = y: seg(Pi) = y
10:      else: seg(NNbest) = seg(Pi) = k, k = k + 1, segk+1 = segk + 1
11:     else
12:       all P with seg(P) = seg(NNbest): seg(P) = seg(Pi)
13:       if exists best NN with seg(NN) = ∅
14:         seg(bestNN∅) = seg(Pi)
15:     else
16:       if seg(Pi) = x: seg(Pi) = x
17:       else: seg(Pi) = segk, k = k + 1 ; segk+1 = segk + 1
18:   end
  
```

Figure 3: Proposed algorithm for façade plane detection

A plane is accepted if it is composed of at least j points, what is again dependent on the point density. As can be seen in the pseudo code the best fitting point has priority. The best fitting point is the point having the minimum product of distance and angle of normal vectors (in radian). Only this point is added to the segment of the processed point, of course only if the both thresholds are not exceeded. If the best fitting point and the processed point are allocated to different segments, these segments are fused and the best point, which has no segment yet, is also added to this segment. For every segment a plane is fitted using principal component analysis performed on the matrix with the points of the plane. The normal vector is the

principal component with the smallest covariance (Klasing et al. 2009). Subsequently the plane is forced to be vertical by projecting the normal vector into the x-y-plane. The planes are intersected to get the vertices of the building. This step is done manually yet, but shall be automated in future.

Not all the points before and behind the façade planes (indoor points, protrusions, intrusions) are included in the segmented façade points because they have not been regarded during the extraction of façade candidates or the regiongrowing process. Because of this there is a step back to the complete point cloud. All points are chosen which are in the range of k m (depends on maximum window height and looking angle of the laser) before and behind the plane. These are finally used for the window detection. From now all façades are processed on their own.

3. WINDOW DETECTION

3.1 Extracting indoor points

First the main façade plane, which is used as reference for the decision if a point lies on or behind it, is determined using a RANSAC algorithm. It is assumed that the plain with the most inliers from the points derived by the bounding box extraction should be the best reference plane.

The points are transformed into a coordinate system with the x-axis being orthogonal to the plane. A histogram is calculated showing the amount of points in a certain distance from the plane (see Figure 4).

Whether a point is declared to be lying behind the main façade plane or not is dependent on the façade roughness (flat surface or many protrusions/intrusions). Because of this a Gaussian function in the following form (see Eq. 1) is fitted to the histogram:

$$y = A \frac{1}{\sigma \sqrt{2 \cdot \pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

The position $\mu - \sigma$ is chosen as threshold for the indoor points (dashed line in Figure 4). The minimum threshold is defined as 10 cm. Finally the indoor points are projected into the façade plane regarding the incidence angle of the laser.

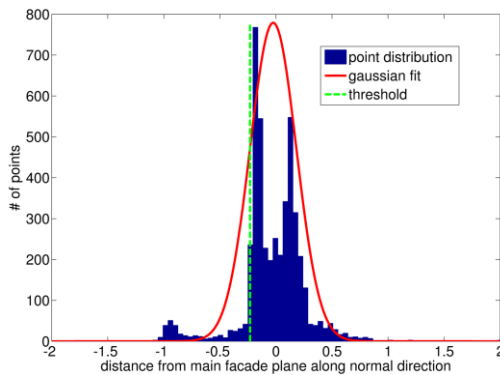


Figure 4. Example for a histogram of the point distribution along the normal direction of the façade, the red line is the fitted Gaussian function and the dashed line is the derived threshold to decide which points are indoor points.

3.2 Binary image and cross-correlation

A binary image with a resolution which is appropriate for the point density (e.g. 1 m) is created from the indoor points by setting a pixel to 1 if a point exists in the respective cell. The binary image is resampled, increasing the resolution with factor 10, and then cross-correlated with a template of a horizontal and a vertical linear mask of 2 m (an average window size) length (see Figure 5). The horizontal line is used to detect the window positions in x-direction (width) and the vertical line to detect the window positions in y-direction (height).

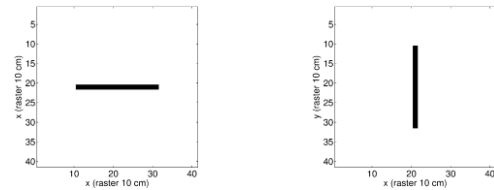


Figure 5. Horizontal and vertical line mask which is used for cross-correlation with the binary image.

Two functions are delineated from the correlation image. The entries of the correlation image are summed up in x-direction to get the positions of windows in y-direction and the other way round. One of these functions can be seen in Figure 6. The use of the line masks works like a smoothing, what leads to clearer peaks in the function of the sums.

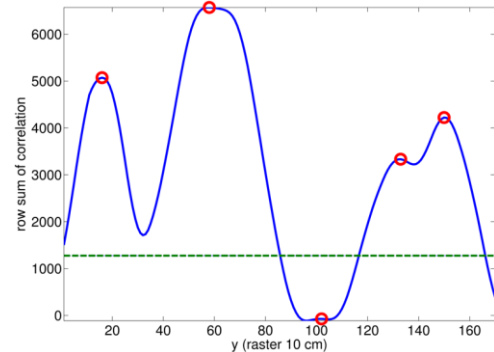


Figure 6. Example for a function of the sum over the values of all columns in the correlation image for each row (= height profile). The peaks in the function, indicating window positions, are marked with circles. The dashed line shows the threshold for accepting peaks as window positions.

3.3 Searching window positions

The search for the window positions is done independently for x- and y-direction. At the end a window is placed at every possible combination of x- and y-positions. The functions of the sums of the correlation images are used twice. First these signals are used as input for a discrete Fourier Transform to look for a repetitive structure. If no such structure can be found the peaks of the function are used as window positions.

As repetition frequency of the windows the best non-zero-frequency of the resulting spectrum is chosen. Three requirements have to be fulfilled to accept the result of the Fourier Transform:

- The frequency has to lead to at least four windows for one façade plane in the respective direction. A signal generated

by fewer windows is too short to provide a reliable solution.

- The median of the distances from the peaks in the function and the frequency have to be equal in between a certain range.
- There has to be a significant peak in the spectrum.

If the result of the Fourier Transform is accepted, the windows are positioned over the whole façade with the determined frequency, starting from the position of the window with the highest correlation (highest peak in the function of the sums). Windows lying too close to the edge of the façade (e.g. <1 m) are neglected.

If the frequency is not accepted, the positions of the peaks in the functions are used. Peaks which are below mean-peak-height/3 (dashed line in Figure 6) are neglected.

The following steps are carried out to improve the derived window positions:

- Peaks which are too close to the edge of the building are removed. This is necessary because there are still some points of the adjacent façades which are normally in front or behind the processed façade, what leads to clear peaks in the function.
- Peaks which are too close together are fused to one peak. The threshold is 1.5 m for the horizontal distance, and 2 m for the vertical distance, assuming that floors have at least a distance of 2 m.

4. EXPERIMENTS

4.1 Data

We use a dataset of the test area TUM (Technische Universität München) recorded by four overflights with a helicopter. The area was scanned in 45° oblique view, what leads to a point cloud, where all building façades can be seen from all

directions. The co-registration of these four different point clouds can be done with homologous planes or an adapted ICP algorithm (Hebel and Stilla 2009, Hebel and Stilla 2007). The composed point cloud can be seen in Figure 7. The total points are around 2.5 million.

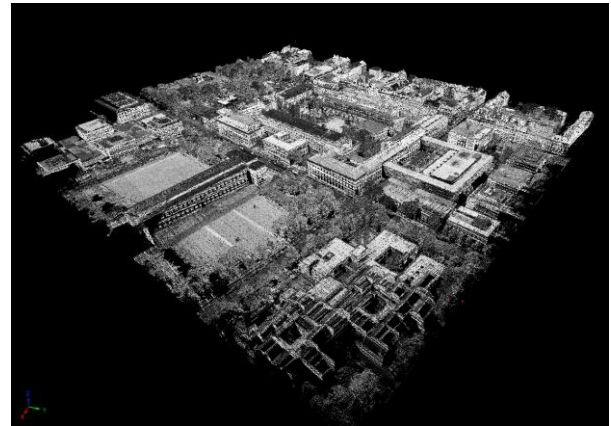


Figure 7. Point cloud with approximately 2.5 million points of TUM area composed of four point clouds from four overflights.

4.2 Façade plane detection

For façade plain detection a threshold of $\Delta\varphi = 10^\circ$ is chosen for the coarse elimination of non-façade-points. Approximately 1/10 of the originally point cloud is remaining.

The algorithm shown in Figure 3 is run with $\Delta d = 3$ m, $\Delta\delta = 5^\circ$ and $n = 10$. To reject small planes the threshold $j = 100$ for minimum points is used. The resulting segments can be seen in Figure 8.

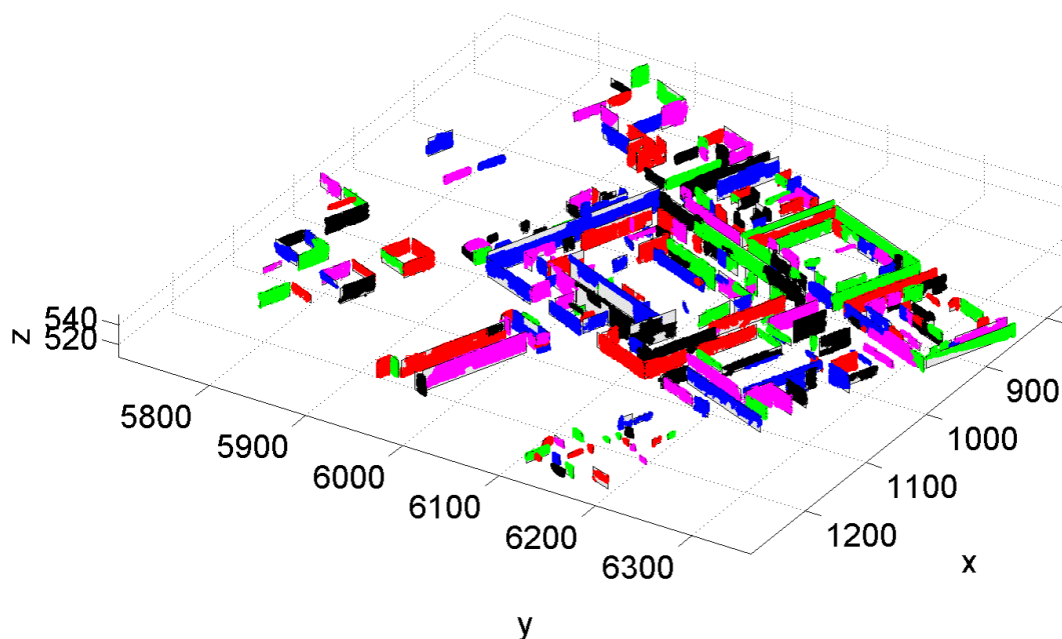


Figure 8. Segments with more than 100 points found by the regiongrowing algorithm

4.3 Window Detection

In the following the window detection is shown for one building (Old Pinakothek) of the test area. The selected points for the northern façade (see Figure 9) received from a bounding box using $k = 5$ m can be seen in Figure 10.



Figure 9. Northern façade of the Old Pinakothek

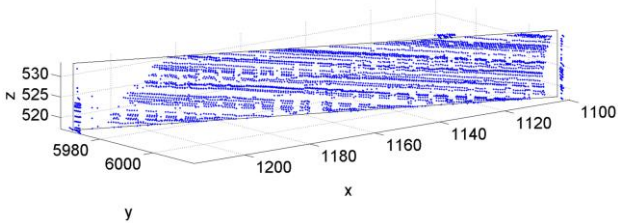


Figure 10. Points derived with a bounding box which delimits the points to the area of 5 m before and behind the façade (northern façade of the Old Pinakothek).

In Figure 11 the detected indoor points are shown. From these points the raster image in Figure 12 is computed.

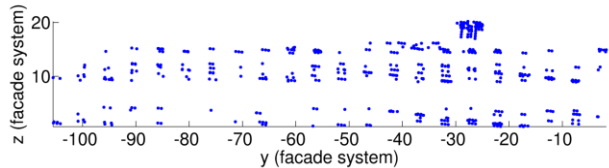


Figure 11. Indoor points derived from the points shown in Figure 10 and the threshold shown in Figure 4.

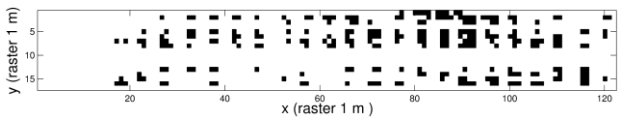


Figure 12. Binary image created by setting a pixel to 1 if any indoor point (shown in Figure 11) is inside the 1 m cell.

Finally in Figure 13 the detected windows can be seen for the façade shown in Figure 9.

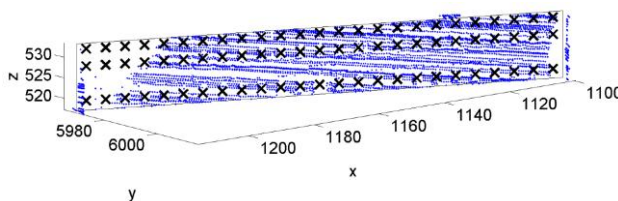


Figure 13. Façade points from Figure 10 representing the façade shown in Figure 9 with detected window (centre) positions.

In this case the number of windows in both directions was determined correctly. Since it is not possible to distinguish between doors and windows, the door in the middle of the façade is also marked as window. An appropriate frequency was

detected for the windows along the width of the façade. For the window positions in height direction the peaks from the function in Figure 6 are used, whereby the last two peaks are fused.

4.4 Results

In Figure 14 ten façade planes of the Old Pinakothek with the detected window centre positions are shown. For evaluation the planes are divided into three groups: a) the two long planes, b) the front face and c) the seven small sides. In Tables 1 to 3 the evaluation results can be seen. Since there are doors in group a) and intrusions in group b) and c), which lie behind the main plane, these are also detected as windows. They are specified separately in the tables.

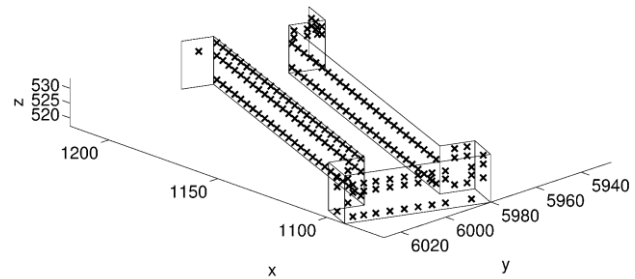


Figure 14. Ten façade planes of the Old Pinakothek, the black crosses represent the detected window positions

Group a)					
W	D	T	WD	DD	FA
123	2	125	123 (100%)	2 (100%)	-

Table 1. Result of the window detection for the two long planes, W = windows (ground truth), D = doors (ground truth), T = amount of detected windows, WD = correctly detected windows, DD = detected windows which are actually doors, FA = false alarm

Group b)					
W	I	T	WD	ID	FA
9	5	18)*	9 (100%)	5 (100%)	4)* 13 (48%)

Table 2. Result of the window detection for the front façade, W = windows (ground truth), I = intrusions (ground truth), T = amount of detected windows, WD = correctly detected windows, ID = detected windows which are actually intrusions, FA = false alarm;)*) these values are produced counting double detection of one (large) window/intrusion as one detection.

Group c)					
W	I	T	WD	ID	FA
7	7	17)*	3 (43%)	7 (100%)	7)* 18 (64%)

Table 3. Result of the window detection for the small façade parts, explanation of abbreviations see Table 2.

As can be seen from the results in Table 1 windows showing a regular pattern over a certain extent, like it is in group a), can be detected very well. For the front plane (b)) the problem occurs that there are intrusions which produce indoor points. These cannot be distinguished from indoor points, which are originated from real windows. Additionally this façade does not show a regular pattern over the whole plane.

The worst results can be found for group c). There is only one window, so the approach searching for a repetitive structure cannot work. The problem with intrusion also occurs here, what leads to many false alarms.

5. CONCLUSIONS AND FUTURE WORK

This paper has shown that a window sequence can be detected, if a signal can be produced by indoor points from a sufficient amount of regular arranged windows. From that follows that this approach is useful for urban scenarios, where often buildings with several floors and regular structure of windows can be found.

The result of the façade detection shows that it is useful to work with multi-aspect side looking airborne laser scanning. From that data façade planes at the back of the buildings, which cannot be seen from the street, or as in the case of the Old Pinakothek, façade planes which are too far away from the street to be acquired by a street mapper, can be provided.

Since the work on this approach is at the beginning, there are several possibilities to improve the results:

- The buildings have to be extracted automatically after the segmentation. This can be done using the normal direction of the façades which can be derived from the navigation data.
- It has to be found a way to distinguish between doors, windows and other intrusions. For that purpose other features like the point density in the façade plane or intensity values have to be considered.
- Separation of rows and columns to make it possible to detect different pattern for different rows/columns of one façade. This is especially important for the ground floor row, which often shows a special behaviour.
- The approach can be extended in a way that also protrusions can be detected.
- The window size shall be derived, for example by fitting a rectangle function in the function of correlation sums.
- The geometric accuracy of the detected window centres has to be evaluated.

ACKNOWLEDGEMENT

The authors want to thank Marcus Hebel (Fraunhofer IOSB) for providing the co-registered point cloud.

REFERENCES

Auer S., Gernhardt S., Bamler R., 2011. Investigations on the nature of Persistent Scatterers based on simulation methods. In: *2011 Joint Urban Remote Sensing Event (JURSE)*, Munich, Germany, pp. 61-64.

Awad T. M., Zhu Q., Du Z., Zhang Y., 2010. An improved segmentation approach for planar surfaces from unstructured 3D point clouds. *The Photogrammetric Record*, 25(129), pp 5-23.

Becker S., 2009. Generation and application of rules for quality dependent façade reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6), pp. 640-653.

Boulaassal, H., Landes, T., Grussenmeyer, P., 2009. Automatic extraction of planar clusters and their contours on building façades recorded by terrestrial laser scanner. *International Journal of Architectural Computing*, 2009 (1), pp. 1-20.

Haala N., Kada M., 2010. An update on automatic 3D building reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(6), pp. 570-580.

Hebel M., Stilla U., 2007. Automatic Registration of laser point clouds of urban areas. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Munich, Germany, Vol. XXXVI, Part 3/W49A, pp. 13-18.

Hebel M., Stilla U., 2009. Automatische Koregistrierung von ALS-Daten aus mehreren Schrägansichten städtischer Quartiere. *PFG Photogrammetrie, Fernerkundung, Geoinformation*, 2009 (3), pp. 261-275.

Iwasczuk D., Hoegner L., Stilla U., 2011. Matching of 3D building models with IR images for texture extraction. In: *2011 Joint Urban Remote Sensing Event (JURSE)*, Munich, Germany, pp. 25-28.

Klasing K., Althoff D., Wollherr D., Buss M., 2009. Comparison of surface normal estimation methods for range sensing applications. In: *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2009, Kobe, Japan, pp. 3206-3211.

Pu S., Vosselman G., 2009. Knowledge based reconstruction of building models from terrestrial laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6), pp. 575-584.

Ripperda N., 2008. Determination of façade attributes for façade reconstruction. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, China, Vol. XXXVII, Part B3A, pp. 285-290.

Schmittwilken J., Plümer L., 2010. Model-based reconstruction and classification of façade parts in 3D point clouds. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Sainte-Mandé, France, Vol. XXXVIII, Part 3A, pp. 269-274.

Vosselman G., Gorte B. G. H., Sithole G., Rabbani T., 2004. Recognising structure in laser scanner point clouds. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Freiburg, Germany, Vol. XXXVI, Part 8/W2, pp. 33-38.