In: Paparoditis N., Pierrot-Deseilligny M., Mallet C., Tournaire O. (Eds), IAPRS, Vol. XXXVIII, Part 3A – Saint-Mandé, France, September 1-3, 2010

MODEL-BASED RECONSTRUCTION AND CLASSIFICATION OF FACADE PARTS IN 3D POINT CLOUDS

Jörg Schmittwilken and Lutz Plümer

Institute of Geodesy and Geoinformation, Bonn University, Germany
{schmittwilken, pluemer}@igg.uni-bonn.de

Commission III - WG III/4

KEY WORDS: lidar, classification, reconstruction, facade parts, 3d city models

ABSTRACT:

The objective of this work is a robust and highly efficient method for reconstruction and classification of facade parts in 3D point clouds. The presented top-down method is based on knowledge derived from a database of ground truth parameters. The usage of prior information is the key feature of our concept. It enables a robust and quick computation that even copes with a high percentage of outliers. Model knowledge is encoded in probability density functions. Furthermore, we combine random sampling with a novel scoring function that rates the size of predictions. Decision trees were applied to identify the most evident features that are assessed by the scoring function. Finally, cluster analysis is employed to estimate those borders of the respective object that cannot be identified precisely by the parametric model. Since the reconstruction of geometric models from a 3D point cloud is close to its classification, both problems are linked up by the presented concept. Our method works on non-oriented point clouds and is free of user interaction. The concept is delineated and demonstrated by the reconstruction of windows and stairs.

1 INTRODUCTION

Terrestrial laser scanning (TLS) rapidly provides highly detailed 3D data. In a few hours, terabytes of 3D point clouds are captured by static or mobile systems from nearly every perspective. TLS provides fast, precise and high resolution measurements. Building facades can be observed with a spatial resolution and deviation less than one centimetre. Mobile, vehicle mounted systems observe more coarse but geo-referenced data of whole streets by linking laser scanners and inertial measurement units.
The interpretation of this data, however, consumes much more time and is, so far, not done automatically. Thus, the construction of highly detailed virtual 3D city models is mainly done manually or semi-automatically yet. The reconstruction of geometrical primitives like planes, cubes, spheres or cylinders is almost solved, but the methods lack semantic interpretation of the reconstructed objects. Clustering methods classify point clouds with the aim of data reduction or reconstruction. But most clustering algorithms also do not consider semantics. On the other hand, just the semantics of objects is important for a lot of analysis done in virtual 3D city models, e.g. coordination of rescue teams or disguising the building interior in TLS because of the right to privacy.
Due to those limitations mentioned above of established methods, i.e. low degree of automation and ignorance of semantics, we present a class of robust model-based classifiers that automatically operate arbitrary 3D point clouds and result in semantical meaningful objects. The demonstrated method combines the classification and reconstruction of TLS data and is exemplarily shown for windows and stairways. In advance, figure 1 shows the result of the reconstruction of windows, door and stair.

In this article TLS data is exclusively subsumed to 3D point clouds and no original observation like time, pulse or waveform information is taken into consideration.

1.1 Related work

Since our approach aims at the classification and reconstruction of meaningful facade objects from 3D point clouds some related topics are reviewed in the following. Clustering and segmentation methods are considered first followed by an overview of actual work on the reconstruction of geometric objects and buildings or building parts.

Figure 1: Estimated windows, door and stair (green) with point cloud (purple) and rectified image

Reconstruction of geometric objects The random sampling consensus (RANSAC) paradigm presented by Fischler and Bolles (1981) is often applied to the robust parameter estimation of geometric objects in point clouds. RANSAC even copes with a high percentage of outliers but is limited to models that are described by a fixed set of parameters. The algorithm proceeds as follows (1) draw a sample of minimal size randomly and instantiate the model, (2) estimate the size of consensus set by counting inliers, i.e. points of which the distance to the model is less than a given threshold ϵ, (3) repeat (1) – (2) until the best model is found. The abortion criterion is statistically founded and dynamically

computed after each iteration. Torr and Zisserman (2000) give a novel scoring function that uses **m**aximum **l**ikelihood **e**stimation. Hence, the modified algorithm is named MLESAC. In contrast to RANSAC it minimizes the sum of residuals and includes a penalty for outliers such that the residual of an inlier is its distance from the model and the residual of an outliers is ϵ.
Vosselman et al. (2004) explore various methods for the segmentation of 3D point clouds received from airborne or terrestrial laser scanning. They conclude that surface growing and 3D Hough transform are the important methods for the reconstruction of buildings and roofs and state that "the most suitable segmentation algorithm may depend on the kind of application".
Schnabel et al. (2007) describe an improvement of the RANSAC paradigm that prefers local sampling, i.e. limiting the single elements of the sample set to a close range. Their concept iteratively estimates several geometric primitives like planes, cylinders or spheres but lacks any semantic information of the reconstructed objects.
Schmitt and Vögtle (2009) present a voxel based method for the extraction of planar surfaces from 3D point clouds. They use normal vectors determined from the 26 neighbors to differentiate points on planar surfaces, edges or corners. They demonstrate their concept by the reconstruction of windows where they assume windows by uninterpreted boundary lines.

Reconstruction of buildings and building parts Boulaassal et al. (2007) segment the point cloud into planar faces by sequential application of RANSAC. They conclude that different objects are fitted by different planes, e.g. main wall, windows or balconies. Boochs et al. (2009) present an approach that uses RANSAC for the estimation of building elements. After the reconstruction of primitive elements they apply an ontology to aggregate elements to objects that are specified in the ontology. Therefore they consider topological and geometrical neighborhood relations.
Pu and Vosselman (2009) give a method that aims at the reconstruction of building models and takes model knowledge that is encoded in vague thresholds into consideration. They claim, for example, that a wall segment is "large" and "vertical" on one hand, and a door segment has a "certain range" and is "on the wall", on the other. Their method detects windows by searching holes, i.e. longest edges, in the Delaunay triangulation of the point cloud. After that, rectangles are fitted into the 3D points on the border of a hole. Moreover their method generates geometrical hypothesis for small occlusions during the estimation of the outline of the facade.
The concept for the reconstruction of buildings that has been suggested by Becker (2009) starts similar to Pu and Vosselman with the detection of holes in facade segments. Windows are modeled as "no data areas" and points on window borders are therefore characterized by having no points in their top, right, bottom or left neighborhood. Afterwards, the point cloud is split in irregular 3D cells defined by the intersection of planes that are given by border points of windows. The point density of each 3D cell is calculated and finally only 3D cells with a point density greater than a given threshold are labeled as window cells.

1.2 Outline of the prediction based approach

In the following we give an overview of the proposed concept for automatic classification and reconstruction of facade parts from 3D point clouds. By *classification* we mean a labeling of data tuples by their estimated class. *Reconstruction* subsumes the derivation of parameters of a given model from a data set, i.e. regression. Since our method estimates location and shape parameters during the classification process and, vice versa, classifies points during the process of parameter estimation, both terms are sometimes used synonymously in this paper.

In contrast to bottom-up or data-driven approaches the presented classifier proceeds in a top-down manner. It operates in three phases: (1) pre-filtering, (2) estimation of the most likely sample and (3) estimation of the remaining boundaries of the respective object. Prior information derived from ground truth data is employed in each phase to reduce computation effort thereby making the procedure more efficient. The prior information is given by probability density functions.

(1) Pre-filtering is applied to single points and reduces the size of the input point cloud considerably. Since our method aims at facade parts the filtering includes the estimation of the wall plane. Dependent on the particular object the filtering is relative to the wall plane and takes the distributions of the object parameters into consideration.

(2) Like RANSAC or MLESAC based approaches our classifier estimates the model that is most likely, however, our method estimates the goodness of a sample in a different way. Since the facade parts like stairs or single windows are considerable small objects according to the whole facade they constitute a tiny subset of the point cloud. Thus the percentage of outliers is huge and RANSAC would be inefficient in calculating the consensus set. Furthermore we keep the size of the sample set small and accept that only a part of the parameters of the model can be estimated. We make use of the spatial character of the underlying models in such a way that their location and shape parameters are partly derived from a sample. Since a given sample constraints the location precisely and the shape parameters approximately we apply these parameters to define a range query to a spatial indexing structure. We call the result set a *prediction* and use the number of the predicted points as an estimator for the goodness of the sample. We maximize the size of the prediction in our algorithm.

(3) Due to general occlusions in the data set, e.g. windowsills, or the requirement of exclusive constraints of a sample, some of the borders of an object cannot be precisely estimated from a sample. Thus, the need for a precise identification of object boundaries, e.g. left and right border of a stair or the top and bottom of a window, is obvious. For that task we apply clustering algorithms that incorporate both, the probabilities of the parameter distributions as well as the features identified by decision trees.

The rest of the paper is structured as follows: In section 2 we give an overview of the assumptions our classifier bases on, the data it operates on, the database of ground truth parameters and the modeling of parameters by probability density functions. Details of the three processing steps of our approach are given in section 3. Section 4 shows results of the processing of real world data of different resolutions and, finally, we draw our conclusion in section 5.

2 DATA, PARAMETERS AND DISTRIBUTIONS

As outlined in the previous section, the presented classifier operates on 3D point clouds from terrestrial laser scanners. It may also work on point clouds of any other origin, e.g. stereo images, as long as they fit the following assumptions. Although, we tested the classifier solely with various TLS data sets (cf. section 4).

The input data sets are assumed to be 3D point clouds that represent building facades. Whereas the facades are not necessarily required to be smooth. The classifier also works well on structured facades that include ornaments, protrusions (oriels or balconies) or any kind of L-, T- or U-shaped ground plots. Moreover, it copes with a high percentage of outliers, e.g. vegetation objects, fences or cars on one hand, and points in the interior of the building that are measured through windows because of transmission of laser beams in glass, on the other hand. In addition the point clouds do not have to be georeferenced and are allowed to be arbitrarily rotated around the vertical (z) axis, though the z-axis of the point cloud has to be parallel to real world vertical axis.Our investigations showed that an exact levelling of the laser scanner meets this demand. Finally, we assume the point cloud to be taken from different view points so that the self occlusion of facade parts is minimized, e.g. there are points observed on both embrasures of a window. However, some parts of the facade that are above the laser scanner are unavoidably occluded, e.g. windowsills.

Since our approach is based on prior knowledge of objects we built up a database of ground truth parameters of facade parts. It contains measurements of all shape parameters (width, depth and height) of 170 stairs, 60 doors, and 560 windows so far and is being extended continuously. Additionally the number of steps of the stairs and the relative location of the doors, i.e. the left, center or right part of the building, is documented. All data is manually taken either from undistorted, rectified and scaled images, or from high resolution 3D point clouds. The first were taken with a Canon 350D or Nikon D700 digital single-lens reflex camera with calibrated lenses. The latter were captured by static scanning with Leica HDS6100 laser scanner which was mounted on the roof of a van to reduce occlusions by cars, fences, or hedges. Figure 2 exemplarily shows the histograms of the parameters of stairs (purple) and windows (green).

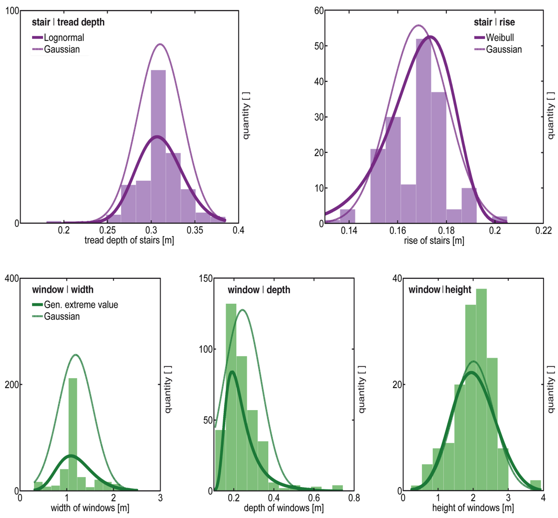

Figure 2: Histograms of ground truth parameters, most likely probability density functions (thick lines) and Gaussian (thin lines) of stairs (purple) and windows (green).

For each parameter that is documented in the ground truth database we estimated the parameters of about 40 probability density functions (PDF) and applied a Chi-Squared test to select the most likely PDF. The distributions of all parameters of windows are optimally described by a Generalized Extreme Values (GEV) distribution, the tread depth of stairs matches a Lognormal distribution and the rise of stairs ties up with a Weibull distribution (cf. figure 2). Apart from rise and depth of windows the Gaussian approximately fits the data nearly as well as the most likely PDF. However, our methods apply the slightly better non-Gaussian probability density functions. The classifier for windows and stairs that are presented in the next section in detail are based on the mentioned distributions.

3 CLASSIFICATION AND RECONSTRUCTION

This section gives a detailed description of our concepts for the classification of facade parts. The following three subsections depict the phases of the classification that have already been outlined in the introduction. The classification and reconstruction of single windows or stairways are delineated. For the detection of multiple objects of the same type, e.g. all windows within a facade, an iterative application of the classifier is required.

Generally, we gradually constrain the point cloud by the application of prior information. Since the constraints are more and more expensive to calculate they are applied to the subset received from the previous step. Therefore, samples are drawn from a subset that is derived from a point based (pre-)filtering. Furthermore the goodness of samples is exclusively computed for samples with a high fitness.

3.1 Pre-filtering

The only restriction on the point cloud that is to be analyzed is the parallel direction of its z-axis and the real world's vertical axis. However, a small amount of uncertainty concerning given directions or angles is permitted. Due to such general assumptions an efficient classifier is needed that copes with a high percentage of outliers. Our method achieves efficiency by its top-down approach that incorporates prior knowledge from the very beginning. Therefore we pre-filter the point cloud with regard to the probability density functions of point distributions to further operate on a subset with a high percentage of inliers.

Our algorithm starts with the estimation of vertical planes by MLESAC. Each estimation operates on the difference set of the original point cloud and the union of the so far received consensus sets. Due to the dominance of walls the estimated planes are assumed to be walls of the building, i.e. (main) parts of facades of L-, T- or U-shaped ground plots or parts of protrusions like oriels. Afterwards, the point cloud is transformed in such a way that the largest plane, i.e. first estimated plane, is identical with the xz-plane of the coordinate system and such that no negative x- and z-values exist. Thus, the positive y-axis points towards the interior of the building. Finally, the reduction of the point cloud is achieved by selecting points with a high probability of belonging to the given class of objects, e.g. points that belong to stairs are most likely on the bottom of the ground floor, in front of the facade and in front of a door. Otherwise points that belong to windows are located on the facade or within a well defined buffer of about 25 cm behind the facade. We apply the probability density functions of model parameters directly or as a mixture of multiple PDFs to the 3D point cloud, and thus receive the filtered subset.

The pre-filtering for windows is solely based on the knowledge about the relative location of embrasures to walls in y direction. The PDF of depth (cf. figure 2) is used to filter points by their y-coordinates relative to the estimated planes.

In contrast to the filtering of windows the filtering of stairs considers three coordinate axis. Hence, the stair filter is a combination of PDFs for x-, y-, and z-coordinates. Each of the three individual PDF is sketched in the following: (x) The PDF of the x value considers the probability of the relative location of the door within the facade and the likelihood of the width of the door. Thus, it is given by a mixture density. (y) The y-filtering is derived as a combination from the probability of tread depth, the PDF of the numbers of steps and the assumption that a stairway is in front of a door and therefore its origin is located at $y = 0$. (z) The PDF of z-location considers the PDF of door heights and the fact that stairs are located at the bottom of the ground floor. The latter is described by a skew-symmetric PDF like Weibull for z-values. Unlike the mixture PDF of x-values the PDFs for y- and z-values are uni modal. Figure 3 shows the pre-filtered points for the estimation of stairs.

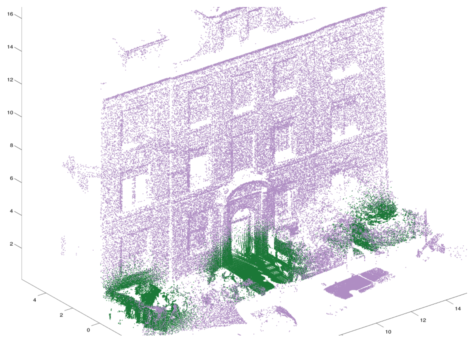

Figure 3: Pre-filtering for the estimation of stairs: Mixture density of x-values and unimodal, skew-symmetric density of y- and z-values.

3.2 Selection of the most likely sample

The sampling is again split into two phases: First, a random sampling and a subsequent fitness proportional sampling and second, the selection of the most likely sample. As mentioned above, the procedure is similar to RANSAC or MLESAC but in contrast to these algorithms we apply a novel scoring criterion for the estimation of the most likely sample. Both, the sampling and the scoring are optimized in order to make the classification more efficient and robust and to reduce computation time.

RANSAC and MLESAC lack efficiency if there is a high percentage of outliers, e.g. a stair that may only come to 0.1% of the whole point cloud. This disadvantage is caused by their sampling strategy: Each sample is supposed to be a valid one (a) and the consensus set is calculated expensively due to a high number of distance calculations (b). We improve the sampling by (a) the rejection of weak samples before calculating the consensus set and (b) introducing a more efficient criterion for the scoring of the goodness of a sample.

More precisely, the rating of a sample means the rating of a subset of its features. The features of a sample may be directly calculable, e.g. the distance of two points, or implicitly given by the neighborhood of the sampled points. The neighborhood of a point is characterized, for example, by the mean of the x-, y- or z-coordinates. Thus, features are properties inherent to the abstract model that describes an object of interest in a parametric way. To find the most evident feature of a model, we calculated the *information gain* (Mitchell, 1997) of each feature. For example, the features of the embrasure model of windows considers the top, right, bottom and left neighborhood of each point. Therefore, the euclidean distance, mean, median, and standard deviation values of the coordinate axis of the four neighborhoods are calculated and ranked by their information gain.

Random and fitness-proportionate sampling Since the knowledge about a model includes its most evident features and the PDFs of its parameter we fit up the sampler with this information and reject poor samples directly after sampling. However, we apply *stochastic universal sampling* (SUS, Baker (1987)) during the rating of the samples. SUS is typically used in genetic algorithms and is a sampling technique similar to roulette wheel selection, but in contrast, it guarantees the expected frequency of the selection of a special individual. The probability that a sample is drawn is proportional to its fitness value. Hence, poor samples are accepted with a low probability and the classification becomes more robust since unlikely hypothesis are also verified to a certain extent. Practically, a number of samples is randomly drawn from the pre-filtered points and their fitness concerning the most evident features is calculated. Afterwards, a smaller subset of the samples is re-sampled proportionally to their fitness by stochastic universal sampling.

In the following we will explain the sampling strategy for stairs. The estimation of stairs subsumes the estimation of treads, i.e. horizontal parts of the stair, and the estimation of risers, i.e. vertical parts of the stair. Here we illustrate the classification of treads. The classification of risers is handled analogously.

Our model of treads considers their *width*, *depth* and vertical distance called *rise*. Thus, one sample consists of two points of which one point is supposed to lie on a tread j and the other one is assumed to lie on the tread after the next tread $j+2$. Treads j and $j+2$ are sampled instead of j and $j+1$ to avoid extrapolation that might be poorly supported, e.g. $j+1$ is the last step of the stair. The hight difference is therefore two times the rise. Here it is obvious that the classification comes along with the reconstruction: The estimation of the stair parameter rise is directly derived from the most likely sample. The sampling for the classification of treads starts with a random sampling of m pairs of points. The $i = 1 \dots m$ probabilities $p_i(width)$, $p_i(treadDepth)$ and $p_i(rise)$ of the coordinate differences $\Delta_{x,i}$, $\Delta_{y,i}$ and $\Delta_{z,i}$ are calculated using the probability density functions. After that, the fitness value of each sample is calculated as the product of the respective probabilities:

$$f_i = p_i(width) * p_i(treadDepth) * p_i(rise).$$

Finally, a subset of $k\%$ of pairs of points is re-sampled by SUS. In the end $m * k$ pairs of points are sampled.

The samples used to estimate windows also consist of two points; one of them is assumed to lie on the right embrasure. The other one is supposed to be on the left embrasure. The fitness of the respective samples considers the horizontal distance in x-direction and the vertical distance of the two points. The "horizontal fitness" ties up with the PDF of the width of windows, whereas the vertical distance is assumed to be small in order to avoid samples of points in different floors of the building.

Scoring by prediction size Since the prior information which is applied within the pre-filtering process is more general we additionally verify the remaining samples with regard to the high-level features identified by decision trees before the calculation of their goodness. If a verification fails the sample is rejected afterwards. Otherwise the goodness of the sample is calculated.

Our scoring function is based on the prediction of the respective instance of the model. The prediction matches the result set of a range query that is defined by the location and shape parameters derived from the sample. Due to the axis parallel transformation of the whole point cloud the spatial query is realized by only one search operation in a kd-tree (Bentley (1975)). Thus, the goodness of a sample with a prediction size a within a point cloud of n points is calculated in $\mathcal{O}(n^{\frac{2}{3}} + a)$.

For example, the query for the prediction of the tread model is specified by the following cuboid of the size (l_x, l_y, l_z) with

$$\begin{aligned} l_x &= median(width_{stair}) \\ l_y &= median(depth_{tread}) \\ l_z &= \sigma_{sensor}. \end{aligned}$$

The cuboid is characterized as "broad and thin". Parallelism of the parameters and the coordinate axis is given by the model assumption of straight stairs that are perpendicular to the facade. Furthermore, the centroid of the range query is located in the mean of the two samples. Figure 4 illustrates the sampling for the classification of stairs.

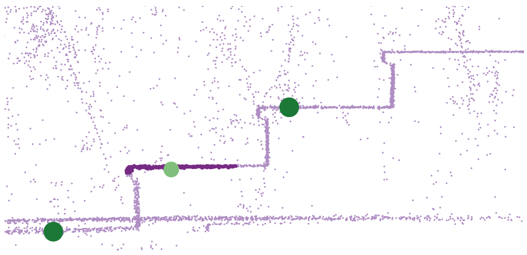

Figure 4: 2D projection: Two sampled points (dark green spheres) and predicted points (dark purple) with centroid (light green sphere) in between.

For windows the constraints of the pre-filtering are not sufficient to ensure that most filtered samples are valid. Therefore, additional verifications are applied to the sample before the prediction of supporters. The following feature was identified by its information gain as the most evident one (out of 40 features) for embrasures: The difference of y-coordinates of the median of two point sets $\mathcal{E}$ and $\mathcal{L}$ is greater than $31cm$. If one of the sampled points is on the left embrasure it defines the centroid of $\mathcal{E}$. The centroid of $\mathcal{L}$ is of the same y- and z-coordinate but translated by $-50cm$ in x-direction, i.e. to the left. $\mathcal{E}$ and $\mathcal{L}$ are received from the kd-tree by a range query with the centroid given by the sampled point and the cuboid.

$$\begin{aligned} l_x &= \sigma_{sensor} \\ l_y &= 2 * max(depth_{window}) \\ l_z &= 0.3 * median(height_{window}). \end{aligned}$$

The factor 0.3 was determined empirically. Thus, $\mathcal{E}$ contains points on embrasures and $\mathcal{L}$ contains points on the facade next to the embrasure. If the sample is accepted $\mathcal{E}$ is the prediction of the left embrasure. The same process is applied to the second sample point, i.e. the right embrasure, analogously.

3.3 Estimation of boundaries

The most likely sample that has been selected during the previous phase of classification only gives a precise description of some parameters of the object. Therefore the boundaries of the object are partly estimated. Windows, for instance, are so far given by one point on each embrasure. Thus, their width is given accurately, however, their height is unknown yet. Since their surface is given by vertical half-planes we have to estimate the top and bottom of the objects. Analogously the width of stairs is unknown so far. Due to the regular shape of man-made objects the boundaries are defined by one reference point and a set of shape parameters. To estimate faces given by rectangular polygons we apply clustering algorithms to the most evident features inferred from the decision trees.

The two sample points of the window model and the prediction of each embrasure, for example, suffice to define thresholds for the x- and y-coordinates. On one hand, the x-coordinate of the median point of the predictions defines the right or left border of the window, whereas the front and the back boundary is approximately given by the minimum and maximum y-coordinate of the predictions. Since no points are observed on windowsills due to occlusion, the estimation of the top and bottom boundary cannot be done analogously to the estimation of left and right embrasure. Therefore it is deduced from the clustering of another evident feature, namely the standard deviation of y-coordinates σ_y along the embrasure in upwards and downwards direction. Similar to the lines in sweep line algorithm (Shamos and Hoey (1976)) we sweep the cuboids of spatial queries iteratively along the vertical line through the points p_j ($j \in \{1,2\}$) of the sample. The starting point and the end are given by

$$\begin{aligned} z_{bottom,j} &= z_j - median(height_{window}) \\ z_{top,j} &= z_j + median(height_{window}). \end{aligned}$$

For each result set r_{ij} the standard deviation of the y-value $\sigma_{y,ij}$ is calculated. The σ_y of cuboids that totally cover the embrasure is much higher than the σ_y of cuboids that cover facades even if they are structured with ornaments. Thus, clustering of σ_y determines the top and bottom of windows. Figure 5 illustrates the sweeping of range queries and the resulting standard deviation of y-coordinates.

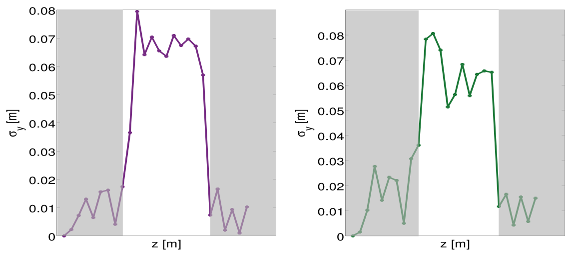

Figure 5: Clustering for the estimation of the boundaries: σ_y of vertically aligned spatial queries of left (purple) and right (green) embrasure. The bright area in the centre indicates the estimated z-range of the window.

4 RESULTS

We implemented the presented methods for straight stairs and windows in MATLAB and applied them to nine 3D point clouds of buildings of the district "Südstadt" in Bonn, Germany. This particular district has been chosen due to its challenging Wilhelmian style buildings of which the facades show sophisticated structures. The data sets that were used for testing were not part of the ground truth database and were not used for the estimation of probability density functions. However, the data sets of which the ground truth database was constructed are also located in the same district. The results are based on the independent estimation of single windows. The estimation was iterated until no more windows were found, i.e. the remaining set of points was smaller than the average number of points of the windows classified so far or a maximum number of iterations was reached. Due to this iterative process the resulting windows have slightly different shape, particularly for the depth.

To optimize the computing time and to show the robustness of our method, we operated on subsets of 100.000 points of the original point clouds. The accuracy of the classification and reconstruction is given in table 1. Three of the 13 *not detected* windows were subsumed by a too large reconstruction that covered two small windows that are close together ("2in1", cf. figure 6 top left and top right windows). Therefore, one of them was counted as not detected, and the other one had a large deviation of width. The reason for the six *erroneously detected* windows was the reconstruction of one ground truth window by two reconstructed windows on top of each other (see figure 6 bottom right window). This typically happened at thick window crosses that were

average point density [pts/m^2]	1135
correctly detected	95
not detected	13
erroneously detected	6
average deviation of x [cm]	1.7
average deviation of y [cm]	2.0
average deviation of z [cm]	19.1
average deviation of width [cm]	5.1
average deviation of depth [cm]	9.8
average deviation of height [cm]	35.5

Table 1: Accuracy of the classification and reconstruction of windows

falsely classified as windowsills or lintels (cf. figure 6). Due to such "2in1" reconstructions, oblique views and occlusions caused by window sills, the *average deviation of z*-coordinates of the reference point of each window, i.e. lower, left, front corner, is ca. ten times worse than its x- and y-values. The high *average deviation of height* is mainly caused by the inaccurate estimation of the reference point. Our method even works for windows with closed shutters (see figure 1).

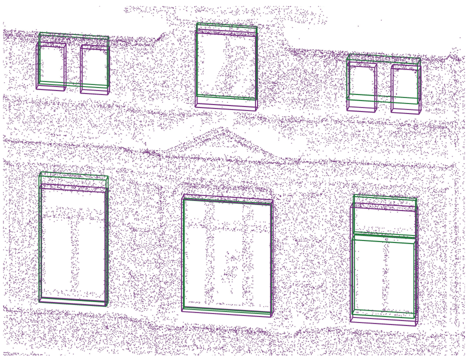

Figure 6: Iteratively and automatically reconstructed windows (green) and manually reconstructed ground truth (purple).

Although the point clouds were taken from the roof of a van, only four of the nine datasets contained non occluded stairs. Their rise and tread depth were estimated with an average deviation of 0.6cm and 2.4cm.

5 CONCLUSION

We introduced a novel concept for the model-based and robust classification of facade parts from 3D point clouds. Prior knowledge was derived from ground truth data and modelled by probability density functions that were applied in many parts of the classification process. The three main steps of the classification algorithm were explained in detail: the pre-sampling, the selection of the most likely sample and the estimation of boundaries. We also presented an efficient scoring criterion for the rating of samples. Therefore the size of the set predicted by a sample was used to estimate the goodness of the sample. The prediction was directly accessible by a spatial indexing structure, namely a kd-tree. Furthermore we applied predictions for a high-level verification of samples. The features underlying this verification were derived from decision trees. We implemented the classifier in MATLAB and applied it to nine non adjusted point clouds of facades observed by terrestrial laser scanning. The convincing accuracy of the automatic estimation of location and shape parameters of the 108 ground truth windows was presented.

We conclude that the application of model-knowledge to top-down classification is feasible and manifold. The representation of prior knowledge by probability density functions and its derivation from ground truth data is sensible. The usage of prior knowledge makes the classification and reconstruction more robust. Furthermore, the rating of samples by their prediction size is highly efficient. The rate of detected objects and the accuracy of the estimation of parameters are satisfactory even for point clouds of low resolution. However, the "2in1 problem" and the lower accuracy of vertical parameters has to be tackled, e.g. by the estimation of window matrices instead of single windows.

References

Baker, J. E., 1987. Reducing bias and inefficiency in the selection algorithm. In: Proceedings of the Second International Conference on Genetic Algorithms on Genetic algorithms and their application, L. Erlbaum Associates Inc., Hillsdale, NJ, USA, pp. 14–21.

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.

Bentley, J. L., 1975. Multidimensional binary search trees used for associative searching. Commun. ACM 18(9), pp. 509–517.

Boochs, F., Kern, F., , Schütze, R. and Marbs, A., 2009. Ansätze zur geometrischen und semantischen modellierung von grossen, unstrukturierten 3d-punktmengen (in german, translated title: 'approaches to a geometric and semantic modelling of large unstructured 3d point sets'. Photogrammetrie, Fernerkundung, Geoinformatik 2009(1), pp. 65–77.

Boulaassal, H., Landes, T., Grussenmeyer, P. and Kurdi, F. T., 2007. Automatic segmentation of building facades using terrestrial laser data. In: P. Rönnholm, H. Hyyppä and J. Hyyppä (eds), International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36, part3/W52, Espoo, September 12-14, 2007, Finland, pp. 65–70.

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

Mitchell, T. M., 1997. Machine Learning. McGraw-Hill.

Pu, S. and 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.

Schmitt, A. and Vögtle, T., 2009. An advanced approach for automatic extraction of planar surfaces and their topology from point clouds. Photogrammetrie - Fernerkundung - Geoinformation 2009(1), pp. 43–52.

Schnabel, R., Wahl, R. and Klein, R., 2007. Efficient ransac for point-cloud shape detection. Computer Graphics Forum 26, pp. 214–226.

Shamos, M. I. and Hoey, D., 1976. Geometric intersection problems. In: Proceedings of 17th Annual Symposium on Foundations of Computer Science, October 25–27, 1976, Houston, TX, USA, pp. 208 –215.

Torr, P. H. S. and Zisserman, A., 2000. Mlesac: a new robust estimator with application to estimating image geometry. Comput. Vis. Image Underst. 78(1), pp. 138–156.

Vosselman, G., Gorte, B., Sithole, G. and Rabbani, T., 2004. Recognising structure in laser scanner point clouds. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36, part8/W2, Freiburg, Germany, October 4-6, pp. 33–38.