EXTRACTING BUILDING FOOTPRINTS FROM 3D POINT CLOUDS USING TERRESTRIAL LASER SCANNING AT STREET LEVEL

Karim Hammoudi, Fadi Dornaika and Nicolas Paparoditis

Université Paris-Est, Institut Géographique National, Laboratoire MATIS
73 Avenue de Paris, 94160 Saint-Mandé Cedex, France
{firstname.lastname}@ign.fr

KEY WORDS: 3D Point Cloud, Hough Transform, RANSAC Method, K-means Clustering, Laser Scanner, Building Footprint, Building Reconstruction, City Modeling.

ABSTRACT:

In this paper, we address the problem of generating building footprints using terrestrial laser scanning from a Mobile Mapping System (MMS). The MMS constitutes a fast and adapted tool to extract precise data for 3D city modeling. Urban environments evolve over time due to human activities and other factors. Buildings are constructed or destroyed and the urban areas are extended. Therefore, the structures of the cities are constantly modified. Currently, building footprints can be generated using aerial data. However, aerial based footprints lack precision due to the nature of the data and to the associated extraction methods. The use of MMS is proposed as an alternative to perform this complex task. In this work, we propose an operational approach for automatic extraction of accurate building footprints. We describe the challenges associated with the terrestrial laser raw data acquired in realistic and dense urban environments. After a filtering stage on the 3D laser cloud point, we extract and reconstruct the dominant facade planes by combining the Hough transform, the k-means clustering algorithm and the RANSAC method. The building footprint is then estimated from these dominant planes. Preliminary experimental results are presented and discussed. The assessments show that this approach is very promising for the automation of building footprints extraction.

1 INTRODUCTION

Nowadays, city modeling has become an important subject of research for architectural lasergrammetry, photogrammetry and computer vision communities. There is an increasing need for 3D building descriptions in urban areas in several fields of application like city planning and virtual tourism. Therefore many research activities on city modeling have focused on the automatic generation of 3D building models from aerial images. Most pipelines which have been developed recover the 3D shape of roof surfaces, but building ground footprints come from existing databases acquired by the digitization and vectorization of cadastral maps or from surveying measurements.

Initially, the building footprints are extracted either in an automatic way using the aerial data (Cheng et al., 2008), (Tarsha Kurdi et al., 2006) or in a manual way requiring many surveyors to make measurements in the terrain. However, these footprint databases sometimes do not exist (e.g., in less developed countries, etc.), can be very difficult to obtain (e.g., in areas with difficult access or prohibited overflights), or can even be of insufficient geometrical quality with respect to some applications. Moreover, the automatic building footprints extraction using aerial images is a hard task. Imprecise and/or incomplete focusing will affect the modeling process in the sense that the final 3D building model will lack accuracy and details.

Recent progress in technologies have allowed the development and the construction of devices for rapid acquisition of 3D cartographic terrestrial data with very high precision in urban environments. The Mobile Mapping System allows an easy coverage of large scale areas such as districts and cities. The feasibility of this kind of system has been demonstrated (Haala et al., 2008), and the usage of this device is increasingly widespread for applications like the conservation of patrimony (Baz et al., 2008) or visualization. Many works using terrestrial laser scanning are particularly focused on segmenting and texturing the building facades (Boulaassal et al., 2007), (Pu, 2008).

This ground-based modeling is thus unavoidable for some applications such as facade texturing where images acquired by a ground based system need to be registered relatively to the aerial 3D model to ensure a satisfactory mapping. Matching the street level images with the 3D aerial model is an extremely complex due to the generalization problems. The data acquired by ground-based 3D data collection systems, can be used to extract and model facades that can advantageously replace the ground footprints in the aerial reconstruction process, thus leading to a coherent use of both aerial and terrestrial data.

This paper focuses on the first step of a global 3D facade reconstruction framework, i.e. the extraction of the facade footprints and planes. The MMS constitutes an alternative and reliable tool which can be useful to obtain building footprints with very high accuracy and details. The aim of this study is to propose an operational approach for automated building footprints extraction in urban environments. The remainder of the paper is organized as follows: Section 2 states the problems related to the raw laser data and their processing. Section 3 presents the proposed approach for extracting the building's footprints and facade planes. Section 4 gives some promising experimental results.

2 OVERVIEW ON PROBLEMS RELATED TO THE LASER RAW DATA AND THEIR PROCESSING

In this study, we use a mobile mapping system for acquiring georeferenced 3D laser point clouds. The Terrestrial Laser Scanning system (TLS system) is a 2D profile scanner. The third dimension is induced by the vehicle displacement. In addition to this, the Mobile Mapping System is equipped with a Global Positioning System (GPS), an Inertial Measurement Unit (IMU) and a Distance Measuring Instrument (DMI), namely an odometer. This equipment was precisely installed by topometry. It allows the gathering of georeferenced 3D laser data with a very high density and also much information about the acquisition (see section 4).

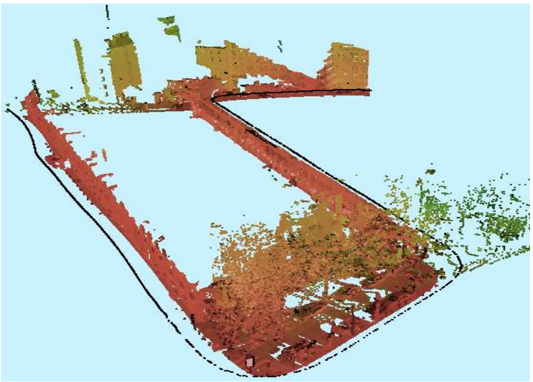


Figure 1: Visualization of the 3D point cloud of very high density. This raw data represents the building facade acquisition. The black line represents the trajectory of the laser sensor.

The laser data are acquired under **realistic conditions** in dense urban environments. Moreover, the datasets are collected in a particular container related to the laser scanner, particularly in a range of 3D points. In this context, the data must be manipulated with much precaution. Here we will describe the main difficulties associated with 3D data.

- **The acquisition:** The laser sensor data can be represented in two ways; either as a container with data organized linearly in temporal sequences of 201 points (data frames) or like a cloud of 3D points (georeferenced data). The conditions of acquisition are very variable in realistic and dense urban environments. Mobile objects cause a thickening of the acquired cloud when the vehicle stops. Certain acquired points model an ephemeral surface and could be considered as erroneous points. Moreover, the density of the facades vary according to the speed of the vehicle.
- **The occlusions:** pose a problem for the complete acquisition of building facades in urban environments. The occlusions could be caused by two categories of obstacles, static or dynamic created by man-made and natural objects. The raw cloud may suffer from missing data due to the presence of pedestrians, trees, mobile and parked vehicles and many others objects (see figure 2). Alas, this inevitable phenomenon affects the modeling process.
- **The laser reflectance:** could cause confusions in the 3D data interpretation. Certain points don't model a physical surface. This effect appears on a retroreflector surface. Observations sometimes show an aureole of points around road signs. These dispatched points represent erroneous data. Moreover, certain points model a different surface other than the surface of interest. Sometimes, the beam of the laser either rebounds off of the outside of the window or it passes through the window and models the inside of the dwelling. These scattered points represent erroneous information for the facades modeling. In addition to this, other less frequent effects could arise due to poorly reflective surfaces.
- **The redundancy of data:** is due to many factors. The acquisition is continuous even when the vehicle is stopped. We have adopted this strategy to facilitate the acquisition of a large area and to use data as common bases for our different projects. Therefore, raw laser data may contain many redundant frames. Moreover, due to sensor characteristics (orientation and linear scanning), we could sometimes have up to three acquisitions of the same facade part caused by the graining of the laser beam in the turns. The redundancy of data (points, frames, parts of the facade) presents an inconvenience for the feature extraction techniques based on vote schemes or random trials.

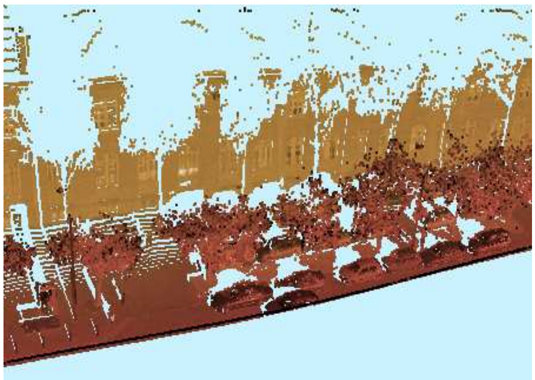


Figure 2: A street in the city of Paris. The building facade is partially occluded by trees and parked vehicles.

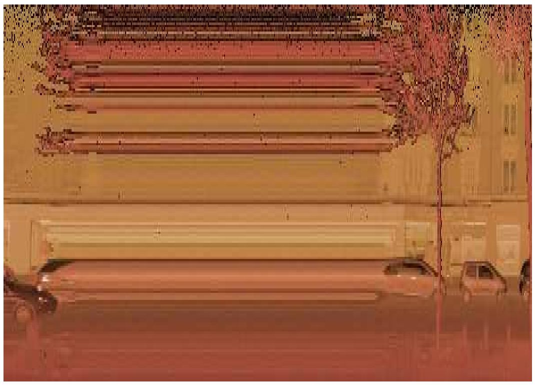


Figure 3: Returned intensities of the 2D scans. The redundancy effect appears when the vehicle temporarily stops. The vehicle and the branches seem stretched.

This brief description allows us to acknowledge several problems associated with the raw laser data. The 3D data should undergo several preprocessing steps before becoming exploitable. Thus we need a process robust to some outliers and noisy data.

3 PROPOSED APPROACH

In this section, we describe our approach which consists of two stages. The first stage focuses on the 3D cloud points preprocessing. The second stage aims at the building footprint extraction. In this work, we assume that buildings have simple polygonal shapes.

3.1 3D data preprocessing

3.1.1 Partial filtering of redundant points As we have mentioned earlier, the laser sensor constantly sweeps the building facade even when the vehicle is stopped. Consequently, the acquired raw data may contain many redundant frames due to this continuous acquisition. For this reason, we have defined a measure between two consecutive frames based on point-to-point distances. The redundant frames are thus detected and removed from

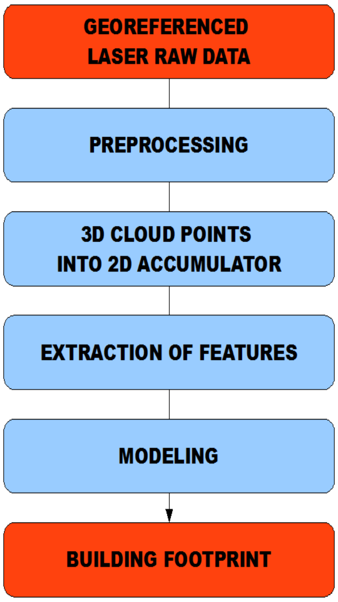


Figure 4: Data flow diagram of the building footprint extraction approach.

the dataset by thresholding these distances. Moreover, this temporal effect distinctly appear on the image of intensity of the laser beam and looks like a succession of rows having the same intensity (see figure 3). Therefore detecting redundant frames can also be based on differences between the intensity of return of two consecutive rows. This step solves only the problem of data redundancy related to the sensor immobility.

3.1.2 Volume of interest The sensor characteristics are used to select the 3D points belonging to the facade plane. The 3D available points are principally positioned above the vehicle to reduce the problem of occlusions. A horizontal band is defined between two horizontal planes. The lower plane passes through the sensor center. The upper plane is shifted by a certain distance that is related to the height of the buildings under study. We precise that the ground altitude could be simply deduced by measuring manually the laser sensor height. Finally, the volume of interest is defined by the georeferenced trajectory of the vehicle and the above horizontal band. The 3D points not included in this volume will be removed from the dataset.

3.1.3 Exploiting the linearity of 3D data After the preceding filtering steps, the frames have undergone a horizontal cropping. The data structure represented by frames is now represented by a sequence of 3D points. We exploit the fact that in this representation facade points are locally aligned. We seek facade points which are principally organized vertically. Thus, the dataset in this sequence is parsed by triplets. The central point of each triplet is kept in the dataset if the triplet is aligned, otherwise it will be removed from the dataset. Therefore, the coplanar points of the building facade are kept. Besides, we observe that the 3D points belonging to other linear structures are also kept.

3.1.4 Mapping the 3D point cloud onto a 2D accumulator The goal of this step is twofold. Firstly, it aims at removing noisy and outlier points. Secondly, it gives a very compact representation of the filtered 3D points. Since we are interested in the vertical structures that generally represent the facades, we project the 3D cloud on a horizontal plane. More precisely, the 3D points are projected into a 2D grid to create an accumulation space. Each point of the cloud votes in one cell, giving a score. Only cells having a high score are kept. The process uses a global threshold which is compared to the maximum score. By this technique, the erratic points of the cloud are removed from the data. The cells with high scores are principally facade points with high density.

Several techniques for the detection of outliers in laser point clouds can be found in (Sotoodeh, 2006) and (Sotoodeh, 2007).

3.2 Building footprints extraction

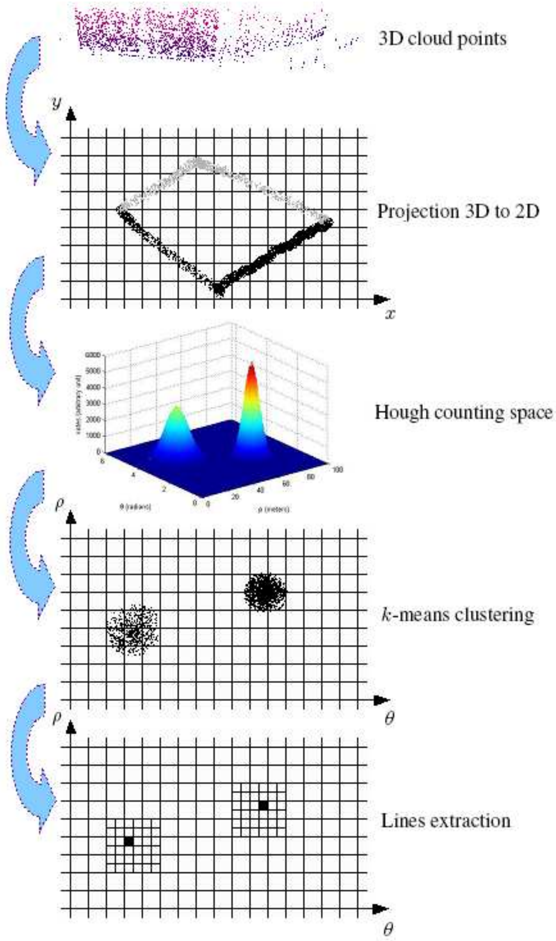


Figure 5: The main steps of our proposed approach for building footprint extraction.

The goal is to automatically extract the building footprint using the 3D filtered points cloud contained in the compact 2D accumulator. The building footprint is a set of 2D segments that can be detected in this 2D space. Recall that the vertical structure of the facades is captured by the scores of the cells. Each cell contains, if any, a set of 3D points $P(x,y,z)$. Furthermore, an efficient extraction can be obtained by working with the barycenters (2D coordinates) of the cells together with their scores. Our approach combines the use of the counting space of Hough Transform, the k-means clustering technique and the RANSAC method. We briefly describe these three techniques and their properties applied in our context. Figure 5 illustrates the main extraction steps.

The Standard Hough Transform (SHT) allows the extraction of the 2D lines among 2D dataset points (Hough, 1962). In the field of our application, this method is currently used to detect the building boundaries in aerial images using the edge points. This method is also used to extract buildings in LIDAR data (e.g., (Tarsha Kurdi et al., 2007) and (Karsli and Kahya, 2008)).

In our approach, we only use the voting steps associated with the Hough transform. We describe briefly here the principle. We made a Hough accumulation space in the discretized parameter space ρ and θ. Each 3D point $P(x,y,z)$ of the dataset (facade points) votes in all cells of the Hough space accumulation verifying the following constraint:

$$\rho = x \cdot \cos\theta + y \cdot \sin\theta \tag{1}$$

where ρ is a length of the normal of the line to the origin and θ is the orientation associated to the normal vector. Each couple (ρ,θ) is unique if $\theta \in [0,2\pi]$ and $\rho \geq 0$.

The cells containing a high score correspond more or less to a facade. However, we aim at determining precisely and automatically the best fit lines of points characterizing the building facades (see figure 6). In our case, we do not carry out the lines extraction step that is based on the local maxima values of the Hough accumulation space since this requires a very difficult tuning of some parameters.

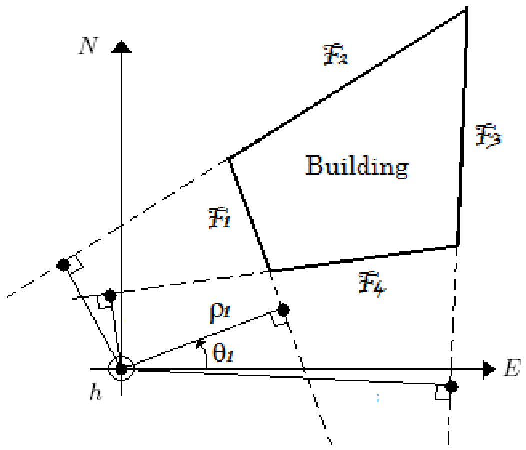


Figure 6: The usage of the Hough transform for extracting building facade lines when the different facades have similar densities.

The local maxima often provide approximate or erroneous solutions. Any partially occluded facade has a low density of 3D points and has thus a low vote in the Hough counting space compared to the non-occluded facade. For this reason, if the vote threshold is too low, many lines will be extracted. Inversely, if the threshold is too high, many lines could be missed. In addition to this, the line estimation depends also on the discretization steps of ρ and θ values. Close lines characterizing the approximation of the same potential line are sometimes extracted. The usage of a large neighborhood to determine the local maxima in the Hough space could reduce this effect. Nevertheless, tuning the threshold is a very difficult task in many cases.

We remind that our approach deals with buildings having simple polygonal footprints. The discretization steps depend also of the building characteristics. The ρ step is related to the minimal distance defined between two facades planes with a similar orientation, the θ step is related to the minimal angle defined between two adjacent facades.

In the urban context, certain characteristics of the building facades are a priori known. The number of facades of the buildings is known in advance (between 3 and 10). Our idea is to use a k-means clustering algorithm to replace the detection of local maxima in the accumulation space—the parameter space of θ and ρ values, in order to automatically determine the exact number of facades and their support 2D lines.

The k-means algorithm is a well-known unsupervised clustering method commonly used to cluster n objects of the input dataset into k homogeneous partitions, $k<n$, for example (Forgy, 1965) and (Macqueen, 1967). We use this technique in a classic way. Nevertheless, several other various clustering techniques exist and a survey is found in (Xu and Wunsch, 2005). Mathematically, the k clusters are determined by minimizing an objective function such as the sum of the squared distances between the points and the corresponding centroids such as:

$$intra\_distance = \sum_{i=1}^{k} \sum_{P_m \in S_i} ((\rho,\theta)_{(P_m)} - (\bar{\rho_i},\bar{\theta_i}))^2 \tag{2}$$

where $(\rho,\theta)_{(P_m)}$ is the value of (ρ,θ) associated with all 3D points $P_m(x_m,y_m,z_m)$ included in the corresponding cell, k is the number of clusters S_i, $i = 1,2,\ldots,k$, and $(\bar{\rho_i},\bar{\theta_i})$ is the centroid of the cluster S_i. The above score is simply the intra-cluster distance measure. More specifically, the score is calculated only for the cells containing a strictly positive vote in the Hough space. Besides, the sum in the above equation is carried out for all 3D point candidates even if they vote all in the same cell.

We can also measure the inter-cluster distance, or the distance between clusters, which we want to be as big as possible. This measure is given by

$$inter\_distance = \min((\bar{\rho_j},\bar{\theta_j}) - (\bar{\rho_i},\bar{\theta_i}))^2, \quad i \neq j \tag{3}$$

Since we want both of these measures to help us determine if we have a good clustering, i.e., a clustering which results in compact clusters which are well separated, we must combine them in some way. The obvious way is to minimize the following objective function:

$$validity = \frac{intra\_distance}{inter\_distance} \tag{4}$$

In our case, the k-means algorithm is run for each k value belonging to the predefined interval. Each run provides a score based on (4). The potential number of facades is the k value corresponding to the minimum of these scores. This validity measure for the determination of the number of clusters in k-means clustering was proposed in (Ray and Turi, 1999). Thus the number of facades could be known even when the facades have heterogeneous densities of 3D points.

More precisely, when one run is carried out for a given k, the algorithm is not guaranteed to return the global optimum because the convergence depends on the initial seeds selected. The k-means algorithm is extremely fast. For this reason, a method which is commonly employed is to run the algorithm several times and select the best clustering available for each k-value. In our case, the first run is carried out by setting the initial seeds to the local maxima in the Hough counting space. The other runs are randomly initialized inside the Hough counting space.

Now that the number of clusters k is known, one can compute a 2D line solution (facade support) for each detected cluster of points. Several solutions can be used to model the facade such as the use of the centroid of each cluster, or the solution with the highest vote for each cluster. We propose to employ a more accurate method such as the RANSAC method to model the building footprint. One advantage of our approach is the following. A curved facade will be approximated by a single line. In addition to this, much information deduced by the clustering step allow to automatically adjust the parameters of the RANSAC algorithm and to thus improve the precision of the detected lines.

The RANSAC method is commonly used to detect lines among edge points (Bretar and Roux, 2005) and (Sester and Neidhart, 2008). We use a classic method (Fischler and Bolles, 1981). For each detected cluster, we use the following process. We use the original space of parameter (x,y). Two different points belonging to the cluster are randomly selected, characterizing a line. A neighborhood is defined along this line by a minimal distance between a point and the line. This process is iteratively repeated until the number of inlier points in the neighborhood is maximized. In our approach, the number of facade points is roughly equal to the number of points for each cluster. Furthermore, the minimal distance associated with the RANSAC technique can be determined from the dispersion of the cluster. When this step is carried out, the best fit lines of 2D points are extracted using the Least Squares Adjustment (LSA technique). A set of 2D segments giving the building footprint is then obtained from the detected 2D lines.

4 PRELIMINARY RESULTS

The acquired 3D data correspond to the facades of buildings in the 12th district of the city of Paris. In this study, we use a high precision 2D laser sensor LMS-Q120i made by RIEGL company [1]. The laser sensor is positioned on the roof of the vehicle. Its beam plane is perpendicular to the vehicle trajectory. The system allows us to carry out 10000 measurements per second and the beam vertically sweeps with an opening of 80 degrees (-20 to 60 degrees with respect to the horizontal). The angular precision of the beam is equal to 0.01 degrees. More specifically, the precision of laser-based measurements is approximately 3 cm at 150 m. In this study, the angular resolution was configured to 201 points by frame (see figure 7). The ground based laser range transmits laser pulses with simple echo.

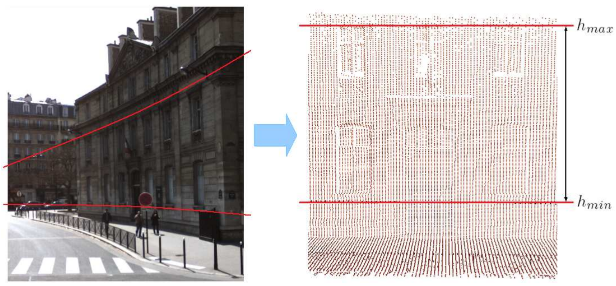


Figure 7: Acquisition of the 3D point cloud using the 2D laser sensor. The frame shows a selected band without occlusions.

The raw measurements provided by the laser sensor are points that are parameterized by distance and angle. The reflected intensity of the laser is between 0 and 1. The coordinates of the 3D points are expressed in the laser sensor coordinate system and also in a common coordinate system, namely the ground reference (absolute) Northern, Eastern and Altitude in Lambert 93. The precision of a 3D point is not easy to evaluate because it depends on the laser precision and on the referencing system precision.

[1]Link to RIEGL company: *http://www.riegl.com*

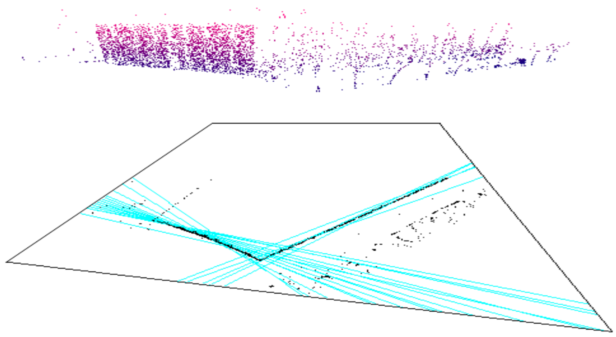


Figure 8: Two difficult facades for the classical Hough Transform: i) a curved facade, and ii) a facade with a low density of 3D points. The detected lines correspond to the local maxima of the Hough space accumulation using the filtered cloud.

The experiments are carried out on two building facades having different architecture and different density of acquired 3D points. One can thus assess the robustness and the efficiency of the proposed approach.

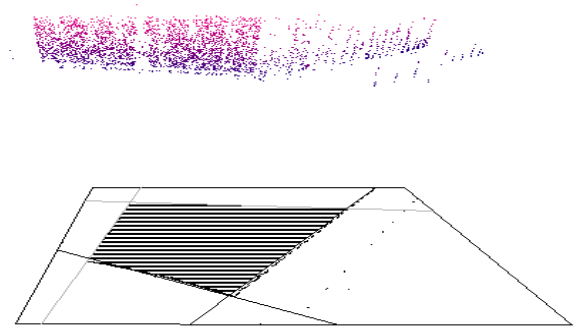


Figure 9: Extracting the building footprint lines using the proposed approach.

Figures 8 and 9 show the extraction of the building footprint lines using the classical Hough Transform and our proposed approach, respectively. The 3D point cloud is presented in the upper part of the figure. The projection onto the 2D accumulator is presented in the lower part of the figure. The studied building illustrates two difficult cases for a classical Hough Transform. Indeed, the left facade does not suffer from occlusions but it is slightly curved. On the other hand, the right facade which is a planar structure is partially occluded, that is, the density of its 3D points in the 2D accumulator space is much lower than that of the left facade.

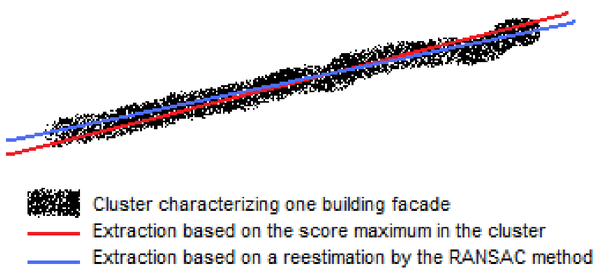
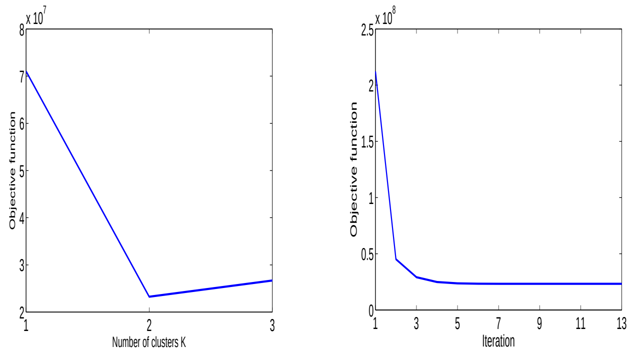


Figure 10: Comparative schema illustrating the precision of lines detection step on one simulated facade.



(a) **(b)**

Figure 11: The application of the k-means clustering.

As can be seen, the classical Hough Transform provided many 2D lines (facade support) corresponding to the many local maxima in the Hough counting space. We can observe that both the curved facade and the partially occluded facade are modelled by several lines. However, by using our proposed approach based on k-means clustering, the correct and accurate 2D lines were obtained. As explained above, the 2D lines can be given either by the centroid of the cluster, its maximum or by the RANSAC technique. As can be seen in figure 10, the building footprint extraction will be more precise using the RANSAC method. The maximum score method detects the line comprising the maximum of points, but it is not necessarily the correct 2D line. The information provided by the clustering method allows us to refine the estimation of the facade lines by exploiting the number of points and the dispersion if the detected cluster (facade) within the RANSAC framework.

Figure 11 shows the application of the k-means clustering algorithm on the 3D data associated with the two facades. Figure 11.**(a)** depicts the validity score as a function of the number of clusters k. As can be seen the optimal value for k is 2. Figure 11.**(b)** shows the convergence associated with this optimum. The footprint lines extracted from this clustering are illustrated in figure 9.

5 CONCLUSIONS AND FUTURE WORK

We presented an approach for the automatic extraction of the building footprint in urban environments. This approach does not require previous knowledge of the number of facades in the input dataset. Moreover, the approach is robust to the heterogeneous densities of facade points. The proposed approach is based on fast filtering and feature extraction techniques. This stage constitutes an essential task for 3D building modeling. Experimental results show the feasibility and robustness of the proposed approach on small islets of buildings.

Future work may investigate the extension of the approach to buildings with a high complexity of shapes and the possibility of application to large areas because each islet of the buildings is delimited by its georeferenced trajectory. Furthermore, since outdoor squares inside the buildings are inaccessible areas for the vehicle, we plan to extend our approach to model full buildings by exploiting the terrestrial data and the corresponding aerial data.

ACKNOWLEDGEMENT

The authors would like to thank Bertrand Cannelle from IGN for his assistance with software and helpful discussions related to the data used in this work.

REFERENCES

Baz, I., Buyuksalih, G., Kersten, T. and Jacobsen, K., 2008. Documentation of istanbul historic peninsula by static and mobile terrestrial laser scanning. In: ISPRS08, p. B5: 993 ff.

Boulaassal, H., Landes, T., Grussenmeyer, P. and Tarsha Kurdi, F., 2007. Automatic segmentation of building facades using terrestrial laser data. In: Laser07, p. 65.

Bretar, F. and Roux, M., 2005. Extraction of 3D planar primitives from raw airborne laser data: a normal driven ransac approach. In: MVA, pp. 452–455.

Cheng, L., Gong, J., Chen, X. and Han, P., 2008. Building boundary extraction from high resolution imagery and lidar data. In: ISPRS08, p. B3b: 693 ff.

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

Forgy, E. W., 1965. Cluster analysis of multivariate data: efficiency vs interpretability of classifications. Biometrics 21, pp. 768–769.

Haala, N., Peter, M., Kremer, J. and Hunter, G., 2008. Mobile lidar mapping for 3D point cloud collection in urban areas: A performance test. In: ISPRS08, p. B5: 1119 ff.

Hough, P., 1962. Method and means for recognizing complex patterns. In: US Patent.

Karsli, F. and Kahya, O., 2008. Building extraction from laser scanning data. In: ISPRS08, p. B3b: 289 ff.

Macqueen, J., 1967. Some methods for classification and analysis of multivariate observations. In: 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA.

Pu, S., 2008. Generating building outlines from terrestrial laser scanning. In: ISPRS08, p. B5: 451 ff.

Ray, S. and Turi, R., 1999. Determination of number of clusters in k-means clustering and application in colour image segmentation. In: Proceedings of the 4th International Conference on Advances in Pattern Recognition and Digital Techniques, pp. 137–143.

Sester, M. and Neidhart, H., 2008. Reconstruction of building ground plans from laser scanner data. In: AGILE08.

Sotoodeh, S., 2006. Outlier detection in laser scanner point clouds. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI-5, pp. 297–302.

Sotoodeh, S., 2007. Hierarchical clustered outlier detection in laser scanner point clouds. In: Laser07, p. 383.

Tarsha Kurdi, F., Landes, T. and Grussenmeyer, P., 2007. Hough-transform and extended ransac algorithms for automatic detection of 3D building roof planes from lidar data. In: Laser07, p. 407.

Tarsha Kurdi, F., Landes, T., Grussenmeyer, P. and Smigiel, E., 2006. New approach for automatic detection of buildings in airborne laser scanner data using first echo only. In: PCV06.

Xu, R. and Wunsch, D., 2005. Survey of clustering algorithms. Neural Networks, IEEE Transactions on 16(3), pp. 645–678.