Semi-automatic Segmentation and Modelling from Point Clouds towards Historical Building Information Modelling

Hélène Macher, Tania Landes, Pierre Grussenmeyer, and Emmanuel Alby

Photogrammetry and Geomatics Group, ICube Laboratory UMR 7357
INSA Strasbourg, France
`{helene.macher,tania.landes,pierre.grussenmeyer,`
`emmanuel.alby}@insa-strabourg.fr`

Abstract. This paper presents a semi-automatic approach for creating a 3D model from point clouds. The proposed approach consists in the development of two successive algorithms. First, the segmentation of the point cloud in geometric primitives is made based on RANSAC paradigm. Then, in a modelling step, the geometric primitives are used for either surface modelling or boundaries extraction and more particularly sectional view extraction. Regarding the results analysis, the developed approach is promising despite some limitations. Not only the limitations, but also potential improvements of our processing chain are discussed. Finally, a bridge towards HBIM is considered.

Keywords: segmentation, modelling, laserscanning, cultural heritage, HBIM.

1 Introduction

Terrestrial laser scanning is a widespread technology which is used to capture information on the geometry of an object in form of point clouds. Laser scanners permit to collect a large amount of accurate data in a very fast way with a high level of details. This technique of 3D measuring is widely used for "as-built" building modelling.

Unfortunately, the creation of 3D models as well as 2D maps or sectional views from point clouds – required usually by archaeologists and architects – remains largely a manual process. Despite some tools spread out in specific software, the generation of these 3D or 2D products remains time consuming and error-prone, especially for historical buildings where complex structures are encountered. Moreover, the result may differ significantly as it depends on the modeler's interpretation of the scene.

The aim of this work is to develop algorithms to automate the segmentation in geometric primitives and the boundaries extraction of point clouds in order to create 3D models.

2 Related Work

Two types of models can be created from point clouds: geometric models which are based on geometric primitives and meshed models. Models are meshed when the

M. Ioannides et al. (Eds.): EuroMed 2014, LNCS 8740, pp. 111–120, 2014.

considered surface is complex or when a high level of details is required. In this case, the result is very faithful to the reality. However, the associated file requires a large storage volume. For models built with geometric primitives, each element is approached by a geometric shape (planes, cones, cylinders, etc.). The file is thus lighter. Hybrid models are also presented by [1]. In this type of model, continuous areas are modeled by primitives and objects details are meshed.

In order to create a geometric model, a previous segmentation of the point cloud into segments characterizing geometric primitives is necessary. Currently, three main methods are used: region growing, Hough transform and RANSAC paradigm ([2] and [3]). RANSAC paradigm is frequently preferred for its robustness and efficiency regarding processing time and noise management ([4]).

A modelling method based on intersection of primitives is proposed by [5]. It can be related to a reconstruction since elements are created in areas where no information is available.

Another model type based on segmented geometric primitives is the wire-frame model. It is produced based on a process of boundaries extractions. The solution widely adopted for outer and inner points extraction after segmentation is a Delaunay triangulation. The contour points are then classified and joined by lines for creating a wire-frame model. This solution is also used by [6] for the semi-automatic extraction of parametric façades models. Contour lines can also be extracted from a prior mesh ([7]). However, for complex buildings, lines extraction based on a previous meshing remains difficult.

3 Developed Approach

Point clouds acquired in three historical sites have been used to test our processing chain. The historical sites, as well as the developed processing chain are presented in this part.

3.1 Test Sites

In order to show the results of the different steps of our approach, the data taken on three different sites are used as input. That's why it's appropriate to briefly present these sites (Fig. 1):

- the Ottmarsheim's abbey-church, built during the 11th century
- the Dugny's church, a fortified church built during the 12th century
- the Châtel-sur-Moselle's fortress build during the 11th century

For these sites, three recording techniques were used: tachometry (single point measuring), photogrammetry and terrestrial laser scanning. Our group is used to merge several techniques in the field of cultural heritage modelling ([8]).

(a) (b) (c)

Fig. 1. Test sites; a) Abbey-church of Ottmarsheim; b) Dugny's church; c) Châtel-sur-Moselle's fortress

3.2 Processing Chain

The developed approach is based on the processing of point clouds exclusively, without any additional information. The processing chain is presented in Fig. 2.

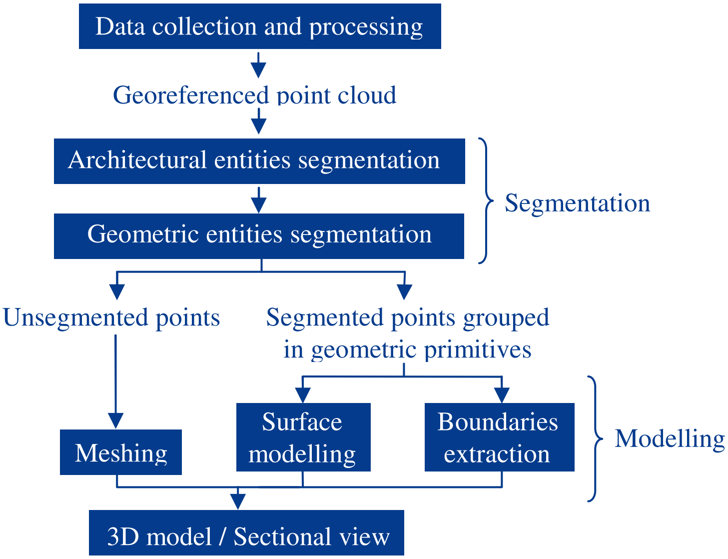

Fig. 2. Processing chain: from point cloud to 3D model

Data Collection and Processing. Registration of the point clouds is made generally with dedicated software. It consists in grouping all point clouds in a unique geodetic system. Then, the obtained point cloud is georeferenced with surveyed points known in coordinates. The data collection and the following registration steps will not be detailed here, because both operations are sufficiently reported in the literature.

Segmentation. The segmentation algorithm aims to extract not only planar segments from point clouds, but also cylinders, cones and spheres. This algorithm, which was developed based on underground spaces of Châtel-sur-Moselle's fortress ([9]), is divided in two main steps: architectural entities segmentation and geometric entities segmentation. As mentioned in the related work, RANSAC algorithm is efficient regarding processing time and noise management. Thus it has been implemented for the extraction of geometric primitives in both steps.

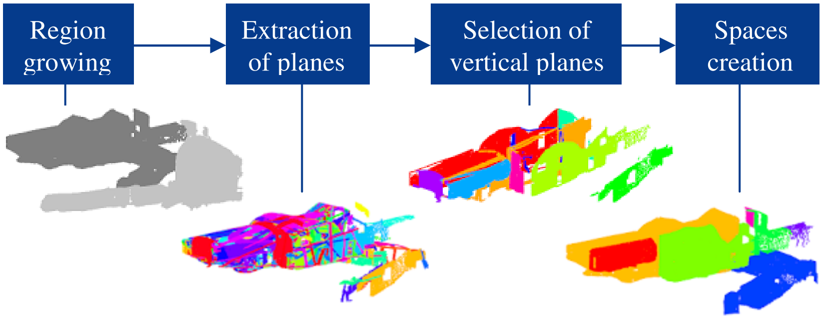

Fig. 3. Architectural entities segmentation

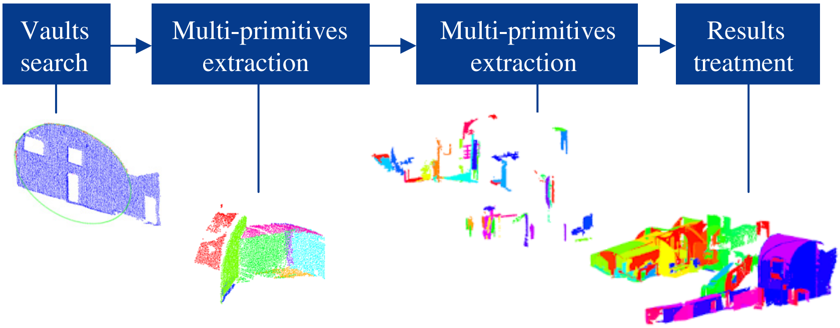

Fig. 4. Geometric entities segmentation

The architectural entities segmentation (Fig. 3) aims to create sub-spaces depending on the main architectural components of the building. First, a region growing algorithm is used to dissociate the elements which are geometrically different. Then, planes are extracted by means of RANSAC and vertical planes are selected with a tolerance of 10 degrees. The major orientation of selected planes is determined and planes following this orientation are considered for the creation of spaces. Planes are selected one by one starting from the smallest plane. For each plane, the closest parallel plane is searched and a space is created.

Then, for each created sub-space, geometric entities segmentation is performed (Fig. 4). First, based on major vertical planes, potential vaults are searched. The method consists in extracting contours of each plane and in searching adjusted circles in these contours. At this stage, parameters of potential vaults are determinate. Multi-primitives including planes, cylinders, cones and spheres are then adjusted and their parameters are stored. After the first extraction, a second extraction is performed with the unsegmented points. Finally, the resulting primitives are sorted based on the number of points associated to primitives, the RMS errors and primitives orientation. The primitives are deleted when their RMS error and the ratio between their superficies and their number of points are too high. Additionally, considering the primitives' orientation, fusions between primitives are carried out.

After the geometric entities segmentation step, the points which are still unsegmented must be analyzed. Either they represent noise or they might provide interesting support to the user for the modelling of details.

Surface Modelling. Once the primitives are detected, the geometric model must be constructed by intersections of primitives. The idea is to consider a plane as a 3D rectangle formed by 4 corners. The goal is to determine the intersections with the four sides if they exist. For each side of a rectangle, several scenarios arise:

- No valid intersection: the line formed by the corner points becomes an edge;
- One valid intersection: corner points are projected onto the intersection line;
- Several valid intersections: the side of the rectangle is decomposed in order to consider the case of only one valid intersection.

Finally, the four lines are intersected to determine the vertices of the plane. This method is not yet implemented and will be the object of further work.

Boundaries Extraction. Based on the previously extracted primitives, boundaries can be extracted. Boundaries are described by contour points which have been extracted through the analysis of meshes based on a Delaunay triangulation. This process of boundary extraction is interesting for creating wire-frame models ([3]) or for extracting sectional views. The semi-automatic creation of sectional view was studied for the Ottmarsheim's abbey-church ([10] and [11]).

The normal vector to the cutting plane defines the view direction of a section. Contours extracted from cylinders (or cones) are dependent on the cutting plane. Apparent contours vary according to view direction. This fact is taken into account in our approach. The cylinder frame is rotated around the cylinder axis until the cutting plane is perpendicular to the view direction.

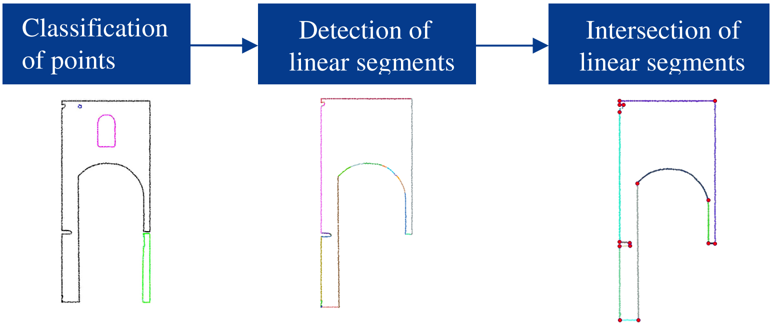

Fig. 5. From contour points to vector contours

The transition from contour points to vector contours is illustrated in Fig. 5. After extracting the contour points of planar and non-planar entities, contour points belonging to the same entity are grouped into one or several classes (Fig. 5 left). A basic line growing algorithm has been developed. A point is selected randomly and its nearest neighbor is searched. If the distance between the two points is lower than a defined threshold (gt), the points are assigned to the same class.

Then, linear segments are detected given a specific threshold (*rt*) with a RANSAC algorithm. Segments which are neither vertical nor horizontal belong to arcs. Using a least square method, the most probable lines and arcs are adjusted to their nearest points. Finally, a wire-frame model can be generated by calculating intersection between curves and lines.

4 Assessment of the Results

At this stage, the segmentation and modelling steps of the developed approach must be evaluated.

4.1 Segmentation Results

Since our method is based on the RANSAC paradigm, a probability is assigned to the segmentation results. For the largest primitives, this doesn't seem to be a problem. But for the smallest, the results may differ slightly from a calculation to another. In order to determine whether this randomness can be problematic, the segmentation results from several calculations of identical primitives were compared. The comparison shows that results are similar. Indeed, for 4 same calculations, the number of points which belong to a primitive varies from 93.2% to 94.8% and the RMS error varies from 2.9 cm to 4.5 cm. The result of segmentation is thus repeatable.

Additionally, a comparison between manual and automatic segmentation was made. The percentage of correctly segmented points was determined using the number of points belonging to manually and automatically segmented primitives, the points in common, the points belonging to automatically segmented primitives that don't belong to manually segmented primitives and conversely. For 9 samples, in average more than 92 % of points are properly segmented.

4.2 Surface Modelling Results

Two tests were made to determine if the extracted primitives correctly describe the geometry of the object.

First, the parameters of automatically extracted primitives are compared with manually extracted primitives. For assessing the quality of planar primitives, the normal of calculated planes is compared to the normal of reference planes. For 10 planar primitives selected randomly, the maximal angle between the normal of the planes is 0.3 degree. This represents a deviation of 5.4 cm for 10 meters, which is more than satisfactory. For one case only, the angle between the calculated and the true normal is larger. It is explained by the fact that the corresponding plane is the smallest one. For 5 randomly selected cylindrical and spherical primitives, the calculated parameters are also close to the reference one, except for one spherical primitive. The unsuccessful modeling comes from the point cloud, which doesn't rightly describe a sphere and this primitive could be classed as false primitive. These comparison shows that parameters are generally close despite some exceptions which could be treated by some filters.

Secondly the extracted primitives were analyzed considering the number of primitives, the percentage of points associated to a primitive and the standard deviation calculated from distances between points and primitives (Table 1). Considering a tolerance of *st* = 5cm, for 95% of points, the distances between points and their associated primitives are less than 5 cm. The same result was obtained for a 3D model created manually in a previous project ([12]).

Table 1. Analysis of segmentation results obtained for 3 different calculations

	Sample 1	**Sample 2**	**Sample 3**
Number of primitives	215	95	49
Number of points	267,711	382,293	70,717
Number of points per primitive	191,851	274,710	66,546
%	71.6	71.9	94.1
Deviation points-primitive I RMS error (cm)	2.9	2.5	3.2
% of deviation < 5 cm	96.37	97.2	96.1

4.3 Boundaries Extraction Results

In order to assess the quality of the automatically created sectional view, a reference sectional view was created manually through digitizing. For this purpose, cutting planes of 6 cm thickness have been extracted from the point cloud and the contours of points slices were digitized in AutoCAD software. The reference sectional view was compared to the sectional view created with our approach.

A quantitative assessment was done by selecting randomly 140 characteristic points between both sectional views. Regarding the results, our approach is promising. Indeed, a mean deviation of 3 cm and a standard deviation of 4 cm have been observed. Moreover, 83% of the deviations are lower than 5 cm.

The lengths of 42 extracted polylines were also compared to the reference ones. Deviations of 2 cm +/- 4 cm have been observed. The extracted contours are obviously located inside the entity, because our approach is based on the points scanned on the object surface. The real edge cannot be scanned, even with very high scanning point density. That's why there are systematically shorter than reference contours.

Our approach is also promising in terms of processing time as it saved 40% of time required for manual drawing, even if manual corrections were done to refine the final model.

4.4 Limitations of the Approach

Several parameters and thresholds have to be defined in our processing chain. They are summarized in Table 2. Obviously, the processing chain is also sensitive to the quality of the point cloud. Indeed, irregular densities or occlusions in the point cloud remain brakes to the automation of 3D model creation from point cloud except if additional information is integrated.

Table 2. Parameters and thresholds involved in the developed processing chain

		Definition	**Selection of parameter value**
Segmentation	*st*	tolerance defining the inliers of the segment using RANSAC	for a point density of 5 cm, *st* was taken equal to 5cm
	w	probability that a randomly chosen point belongs to the best primitive	*w* decreases when the number of primitives increases; generally *w* = 0.2 to 0.5
Boundaries extraction	*ds*	threshold for contour points extractions	sensitive to point density and the level of details to extract; *ds* = 5 cm to 12 cm for point density of 1.5 cm
	gt	maximum distance between two successive points of the same polygon	can be taken as equal to *ds*
	rt	threshold for best lines extraction using RANSAC	depends on the point density and on the quality of the contours; 2 cm for point density of 1.5 cm and *ds* = 5cm

4.5 Potential Improvements

The developed approach uses exclusively the spatial information provided by the point cloud, namely the X, Y and Z coordinates. It would be interesting to integrate the point cloud color and intensity to our processing chain, especially for improving the segmentation results. Indeed, points belonging to a same geometric primitive will presumably have similar RGB and intensity information. Additionally, radiometric changes might be a support for detecting limits between primitives. Moreover, the photographs taken simultaneously to the point cloud might constitute useful information for users to check the model or to complete complex parts of the building which were not or badly modeled.

5 Towards HBIM

Building Information Models (BIM) were originally used for conception of new buildings, but they are also increasingly used for existing buildings. When it concerns historical buildings, the term of Historical Building Information Model (HBIM) is used.

In this type of models, objects are defined in terms of building elements such as walls, roofs, floors, columns, beams, etc (semantic) and described through their geometry. The parameters of objects are stored in a database joined with the 3D model. A BIM is thus not only a 3D model as it combines a 3D model and a database.

An HBIM is particularly helpful for restoration, documentation and maintenance of historical buildings. According to [13], the need to refurbish the cultural heritage is becoming more important than the construction of new buildings.

In order to create an HBIM, as-is or as-built condition has to be modeled. For existing buildings, documentation is often inexistent or not up-to-date. In this situation, laser scanning is used to measure the existing conditions of buildings. As it was mentioned earlier, the process of modelling the captured point cloud is mostly manual and needs to be automated.

The automatic reconstruction of as-built building information models from point clouds is studied in [14]. It involves three tasks: modelling the geometry of the elements, assigning an object category and material properties to an element, and establishing relationships between elements.

Regarding the segmentation and modelling algorithms described in previous sections, our approach allows the geometric modelling of building's elements. Currently, semantic information is not yet considered. It must be included in our approach for orienting the created model towards HBIM.

In each sub-space created with architectural entities segmentation a category of object could be assigned to primitives which define major components of the space. For example considering the position, orientation and size of planar primitives, these primitives could be assigned to four categories: wall, floor, ceiling or none of these categories.

In addition, the potential improvements exposed in the latter paragraph, namely the support of RGB and intensity information, could be used not only to obtain a better segmentation but also to collect information about material properties. Indeed, the mean color and intensity of points of a given primitive could permit to differentiate the type of material such as wood, metal or stone.

6 Conclusion and Future Work

The approach presented in this paper is based on two algorithms and allows to generate 3D and 2D products from point clouds. The results of this approach have been shown using the data acquired in three historical sites. The proposed approach provides satisfying results concerning for instance the processing time and the primitives' adjustment.

However, the degree of automation must be improved. At specific processing steps thresholds are required. Their values depend on the characteristics of the edifice and the scanning parameters.

In the future, the approach will be improved and tested on other historical sites. Other information will be introduce in our processing chain namely RGB and intensity information. Finally, the creation of HBIM can also be developed based on sectional views.

Acknowledgments. The authors want to thank for their support the municipality of Ottmarsheim, as well as the community of communes of Val de Meuse, the association of Vieux Châtel and the French College of Chartered Surveyors. They also want to thank the master students S. Bidino, D. Klinghammer involved in the algorithms development.

References

1. Rusu, R.B., Blodow, N., Marton, Z.C., Beetz, M.: Close-range scene segmentation and reconstruction of 3d point cloud maps for mobile manipulation in domestic environments. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1–6 (2009)
2. Pu, S., Vosselman, G.: Automatic extraction of building features from terrestrial laser scanning. In: Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 36(5), pp. 25–27 (2006)
3. Boulaassal, H., Landes, T., Grussenmeyer, P.: Automatic extraction of planar clusters and their contours on building façades recorded by terrestrial laser scanner. International Journal of Architectural Computing 7, 1–20 (2009)
4. Tarsha-Kurdi, F., Landes, T., Grussenmeyer, P.: Hough-transform and extended RANSAC algorithms for automatic detection of 3D building roof planes from LIDAR data. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Systems, pp. 407–412 (2007)
5. Schnabel, R.: Efficient Point-Cloud Processing with Primitive Shapes. PhD thesis, Universitäts und Landesbibliothek Bonn (2010)
6. Boulaassal, H., Chevrier, C., Landes, T.: From Laser Data to Parametric Models: Towards an Automatic Method for Building Façade Modelling. In: Ioannides, M., Fellner, D., Georgopoulos, A., Hadjimitsis, D.G. (eds.) EuroMed 2010. LNCS, vol. 6436, pp. 42–55. Springer, Heidelberg (2010)
7. Weber, C., Hahmann, S., Hagen, H.: Sharp Feature Detection in Point Clouds. In: IEEE International Conference on Shape Modeling and Applications (SMI), pp. 175–186 (2010)
8. Grussenmeyer, P., Alby, E., Landes, T., Koehl, M., Guillemin, S., Hullo, J.-F., Assali, P., Smigiel, E.: Recording approach of heritage sites based on merging point clouds from high resolution photogrammetry and terrestrial laser scanning. In: Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, pp. 553–558 (2012)
9. Klinghammer, D.: Génération de modèles 3D simples d'espaces bâtis par extraction de primitives géométriques. Master Thesis of INSA Strasbourg (2013)
10. Bidino, S.: Etude de l'extraction automatique de coupe à partir de nuages de points. Master Thesis of INSA Strasbourg (2013)
11. Landes, T., Bidino, S., Guild, R.: Semi-Automatic extraction of sectional view from point clouds. The Case of Ottmarsheim's abbey-church. In: Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Italy, June 23-25 (2014)
12. Alby, E., Grussenmeyer, P.: From Point Cloud to 3d Model, Modelling Methods Based on Architectural Knowledge Applied to Fortress of Châtel-sur-Moselle. In: Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, pp. 75–80 (2012)
13. Giudice, M.D., Osello, A.: BIM for Cultural Heritage. In: Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, pp. 225–229 (2013)
14. Tang, P., Huber, D., Akinci, B., Lipman, R., Lytle, A.: Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. Automation in Construction 19, 829–843 (2010)