Cluster Analysis and Priority Sorting in Huge Point Clouds for Building Reconstruction

Wolfgang von Hansen Eckart Michaelsen Ulrich Thönnessen
FGAN-FOM, Gutleuthausstr. 1, Ettlingen, Germany
E-mail: wvhansen@fom.fgan.de

Abstract

Terrestrial laser scanners produce point clouds with a huge number of points within a very limited surrounding. In built-up areas, many of the man-made objects are dominated by planar surfaces. We introduce a RANSAC based preprocessing technique that transforms the irregular point cloud into a set of locally delimited surface patches in order to reduce the amount of data and to achieve a higher level of abstraction. In a second step, the resulting patches are grouped to large planes while ignoring small and irrelevant structures. The approach is tested with a dataset of a built-up area which is described very well needing only a small number of geometric primitives. The grouping emphasizes man-made structures and could be used as a preclassification.

1. Introduction

In addition to photogrammetric imaging, laser scanning has become a major data source for the acquisition of 3D city models for tourist information or the documentation of cultural heritage. Airborne systems are widely used but also terrestrial laser scanners are increasingly available. They provide a much higher geometrical resolution and accuracy (mm vs. dm) and are able to acquire building facade details which is a requirement for realistic virtual worlds. However, the operating area is limited due to obstruction and by the maximum range of the laser beam. It is not possible to capture an extended area from one position alone, leading to typical raw datasets consisting of several overlapping, huge point clouds. The overall objectives are to

- generalize the point cloud to higher level geometric primitives and group them to objects for data reduction and as preparation for high level cognitive tasks.
- fuse the datasets and coregister them into a single geometric reference frame using these primitives.

While automatic matching of multiple point clouds still is a topic of research [2, 6], different approaches for segmentation and representation through geometric primitives exist. Segmentation techniques include clustering based on local surface normal analysis [1, 6], region growing using scan geometry and point neighborhoods [2], generation of a discrete grid [8], or a split-and-merge scheme applying an octree structure [9]. The objects are often represented by planar elements recovered through RANSAC schemes [1], creation of a triangular irregular network [4], tensor voting [8] or least squares adjustment [9]. The tensor voting scheme is able to describe not only planes, but also linear structures like high-voltage lines.

In this paper, a preprocessing technique is presented, that transforms the point cloud into a set of object surfaces via a two stage process. The first stage is a RANSAC based generation of locally delimited surface patches from the point cloud. The second stage groups patches belonging to the same surface in object space.

2. Generation of surface patch elements

The input is a cloud of 3D measurements gathered by a terrestrial laser scanner and locally delimited planes shall be extracted as surface patches. These may represent a part of larger, planar objects but may as well coincide with small object surfaces. The transformation is split up into two subprocesses, a partitioning of the point cloud into spatial bins and the robust estimation of the dominant plane in each bin.

2.1. Spatial data partitioning

The set of 3D points is partitioned and assigned to 3D volume cells using a Cartesian raster. All points in one of the raster cells will be denoted by $\mathcal{X}$. Two objectives are satisfied by this partitioning:

- The raw point cloud is organized into small blocks that can be processed seperately.
- Since each cell only covers a small part of the complete scene, it emphasizes local features in object space.

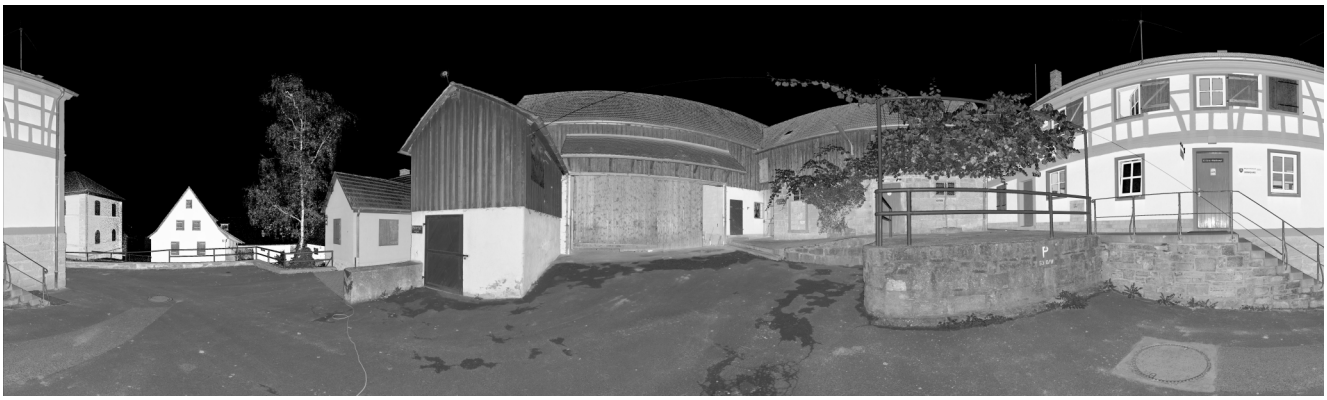

Figure 1. Panoramic image of the test area showing the amount of reflected light as grayvalues.

Since the laser scanning device is set up roughly aligned to the horizontal plane, there exists an implicit alignment between cells and the world coordinate system.

We have to deal with a quantization problem: On the one hand the raster should be chosen as small that any important object surface dominates at least one cell. On the other hand, the cells should not be too small because unwanted microstructure would dominate the results. For the experiments, a raster of 1 m was chosen in order to describe facade and roof surfaces while safely ignoring small structures.

2.2. Robust estimation of plane patches

The second stage is the robust estimation of the dominant plane – the one that has the biggest support from the 3D points $\mathcal{X}$ – independently for each of the raster cells. Only one plane is computed for each cell which seems counterintuitive, but if the cell size is chosen such that almost every significant object surface will be the dominant plane in at least one cell, no information about the object is lost.

Robust estimation implies a result free of outliers – in our case this involves not only the estimation of the plane parameters $\mathbf{p} := (\mathbf{n}, d)$ (Eq. 2), but also of the set of points $\hat{\mathcal{X}} \subseteq \mathcal{X}$ lying on the plane. We have implemented it using the well known RANSAC strategy [3]: A minimum set of three non colinear points $\mathbf{X}_i \subseteq \mathcal{X}$ is chosen randomly and the uniquely defined plane $\mathbf{p}_i$ is computed. Then for all points $\mathbf{x}_j$, their distance d_{ji} to the plane $\mathbf{p}_i$ is computed, defining the set of inliers of run i as

$$\hat{\mathcal{X}}_i := \{\mathbf{x}_j \in \mathcal{X} \mid d_{ji} \leq d_{\max}, j = 1 \ldots |\mathcal{X}|\} \qquad (1)$$

These steps are repeated, while the largest inlier set $\hat{\mathcal{X}}$ is retained as result. In a final step, the final plane parameters $\hat{\mathbf{p}}$ are estimated from all points in $\hat{\mathcal{X}}$.

The plane represented by the Hesse normal form

$$ax + by + cz + d = \mathbf{n}^\top \mathbf{x} + d = 0 \qquad (2)$$

has an infinite extent. We are interested in a small and delimited plane representing the points $\hat{\mathcal{X}}$ only. Therefore, in addition to the normal vector $\mathbf{n}$ and the distance d to the origin, the mean $\mathbf{o}$ of the point cloud $\hat{\mathcal{X}}$ is stored as well.

For the visualization, all points $\hat{\mathcal{X}}$ are projected orthogonally onto $\hat{\mathbf{p}}$ thus forming a 2D point set $\hat{\mathbf{x}}$. The convex hull of $\hat{\mathbf{x}}$ is determined and its (planar) polygon transformed back into 3D space as the boundary of the plane element. Similar to the 3D data partitioning, the 2D points are partitioned and assigned to a 2D raster in order to compute a texture image. Some of the 2D cells may remain empty because the point cloud is sparse or contains holes. A filling factor λ is computed and will be used as weight for the grouping step. Subsequently, such planes are called patch and denoted by $\mathbf{P}$.

3. Grouping of planes

Each patch instance $\mathbf{P}$ has the attributes $(\mathbf{n}, d, \mathbf{o})$ assigned with it. Regarding coplanarity as a major property of man-made objects, we group these instances into planar cluster objects $\mathbf{C}$

$$\mathbf{C} \in \mathcal{P}(\mathbf{P}) \wedge \forall \mathbf{P}_i, \mathbf{P}_j \in \mathbf{C} : \mathbf{n}_i^\top \mathbf{o}_j \approx d_i, \qquad (3)$$

where $\mathcal{P}(\mathbf{P})$ is the power set of $\mathbf{P}$ and $\approx$ means approximately equal. Given such an object instance $\mathbf{C}$ we again assign attributes $(\hat{\mathbf{n}}, \hat{d}, \hat{\mathbf{o}})$ to it: $\hat{\mathbf{o}}$ by averaging over $\{\mathbf{o}_i\}$, $\hat{\mathbf{n}}$ as null space of the set $\{\mathbf{o}_i - \hat{\mathbf{o}}\}$ and $\hat{d} = \hat{\mathbf{o}}^\top \hat{\mathbf{n}}$.

The set of all objects $\{\mathbf{C}\}$ is a subset of $\mathcal{P}(\mathbf{P})$. We propose to list only the few most important in a priority ordered sequence. First criterion for the importance of an object $\mathbf{C}$ is its size $|\mathbf{C}|$. Second criterion is its 2D property captured by the ratio between the third and the second singular value s_3/s_2 acquired during null space determination. As combined criterion we choose

$$c := -\ln(s_3/s_2) \cdot |C|. \qquad (4)$$

Figure 2. 3D model of the main building.

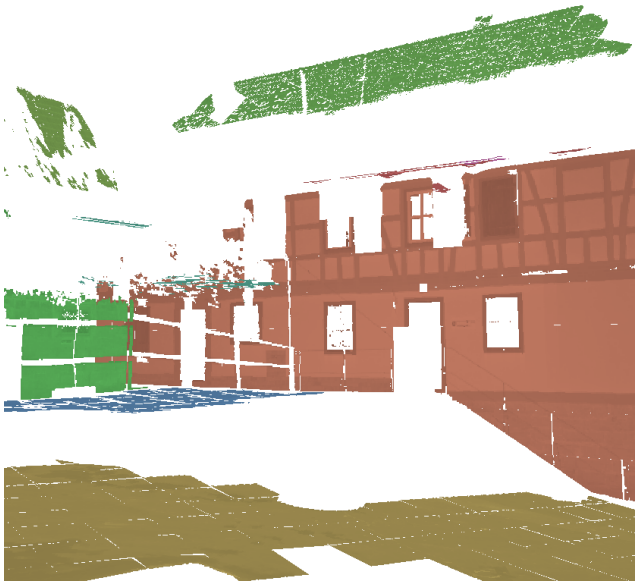

Figure 3. Remaining major planes from Fig. 2.

The search for an approximation of this ordered list of objects $\mathbf{C}$ is fostered by a sensible assessment criterion on the objects $\mathbf{P}$ for which we utilize the filling factor λ (Sec. 2.2). The list of all patch elements $\mathbf{P}$ is prioritized according to λ. The production of cluster objects is started by picking *seed* objects with the highest priority. For each such object $\mathbf{P}_i$ the set of possible partners $\{\mathbf{P}_j\}$ is queried for

$$||\mathbf{n}_i - \mathbf{n}_j|| < \mathbf{t}_n \wedge |\mathbf{n}_i^\top \mathbf{o}_j - d_j| < t_0. \qquad (5)$$

The maximal set of objects $\{\mathbf{P}_j\}$ fulfilling this aggregating constraint is used for construction of a new object $\mathbf{C}$.

We emphasize that this procedure does not exactly give the set of cluster objects defined by Eq. 3: It suppresses subsets of bigger clusters, and it does not really test for $\mathbf{n}_j^\top \mathbf{o}_i \approx d_i$. The similarity condition for $\mathbf{n}_j$ and $\mathbf{n}_i$ is used for faster query handling and the simple threshold t_0 for $\approx$ in the coplanarity query is a rough approximation being aware of work such as [5]. For t_0 we used 6 cm here, which is double of what has been used in Eq. 1. The approach has a bias listing more dominant and important objects first, so that the search can be interrupted by any external demand and then output the set of objects $\mathbf{C}$ obtained so far.

There will be many similar objects in the output set, differing in very few predecessors only, due to straight forward control structure. This motivates a second ordering procedure operating on the output set of the first. It computes the criterion c_i (Eq. 4) for each object $\mathbf{C}_i$ and picks the object with maximal c first. From this object the attribute vector $(\mathbf{n}_i, d_i, \mathbf{o}_i)$ is taken to compute a suppressing factor

$$f_{ij} = 1 - \mathrm{e}^{-||\mathbf{n}_i - \mathbf{n}_j||^2 - \lambda|d_i - d_j|^2 - \gamma||\mathbf{o}_i - \mathbf{o}_j||^2}. \qquad (6)$$

Before picking the next element this factor is multiplied to all the assessments c_j resulting in suppression of objects similar to the first one.

We see close relationship to perceptual grouping discussed in the computer vision community. [7] Coplanarity can be viewed as straightforward generalization of colinearity to just one dimension more.

4. Experiments and results

The investigated dataset has been acquired by a terrestrial laser scanner Z+F Imager 5003. The chosen scene (Fig. 1) is a courtyard surrounded by several buildings at various distances. The reflected light from the laser beam is recorded as a grayvalue and is used for texturing. Because no light is returned from the open sky, it appears black. This dataset contains about 100 million 3D points and covers a spherical area with a radius of 50 m.

The extracted planes are shown for two different viewpoints in the scene (Figs. 2/4). Information is missing at some object plane boundaries because of the restriction of only one plane per raster cell. Parts of the roof structure in Fig. 2 are missing due to obstruction through eaves gutter and pergola (cf. Fig. 1). The empty circular area on the ground gives a good hint for the position of the scanner. The tree (Fig. 4) has been modeled quite well, even though the approach demands a description through planes only. The grouping lead to 24 object surfaces shown in different colors (Figs. 3/5). Only man-made structures are left over while the tree and other smaller objects have been successfully removed.

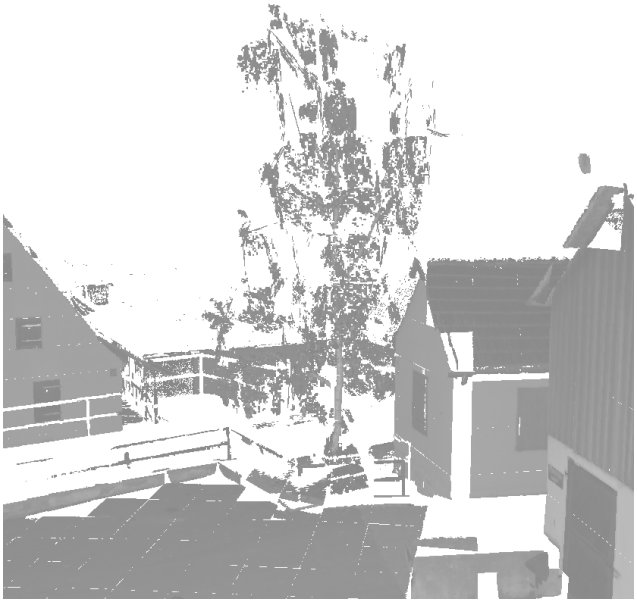

Figure 4. Example of a natural object (tree).

Table 1. Data amount at different stages.

	#points / 10^6	#planes	#surfaces
Point cloud	104	—	—
Small planes	1.7	2616	—
Object surfaces	0.8	1103	24

Tab. 1 shows the data reduction through the stages. The number of points is reduced to 2% through averaging the raw points over texture pixels of a size of 3×3 cm. The complete scene is described with only 2616 planar elements instead of over 100 million points. The grouping step reduces the data again to 1% so that only 24 object surfaces remain. The selection involved with the grouping rejects about half of the small planes as unimportant.

5. Conclusions and future work

We have shown a robust preprocessing method to extract planar surfaces from 3D point clouds. The visualization of this intermediate data shows a good coverage of the scene. The grouping approach selected man-made structures very well so that this step can already be regarded as a preclassification. Giving a brief ordered priority list of important probably man-made surfaces of a scene means an important step towards automatically understanding the scene.

Future tasks are registration of multiple datasets based on the extracted planes, solving conflicts and refinement of the grouping step and a grouping of surfaces to objects.

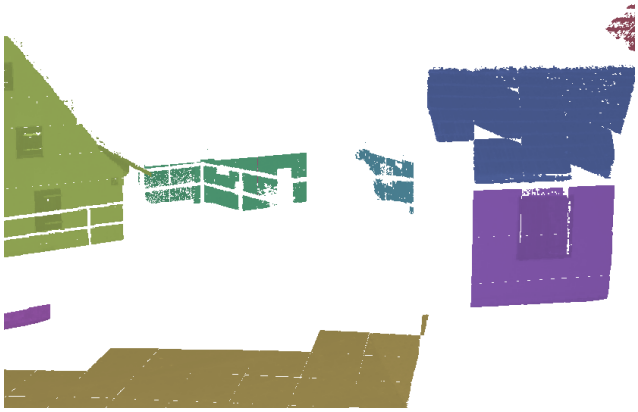

Figure 5. Remaining major planes from Fig. 4.

References

[1] F. Bretar and M. Roux. Hybrid Image Segmentation Using LIDAR 3D Planar Primitives. In G. Vosselman and C. Brenner, editors, *Laser scanning 2005*, volume XXXVI-3/W19 of *IAPRS*, 2005. URL: www.commission3.isprs.org/laserscanning2005/papers/072.pdf.

[2] C. Dold and C. Brenner. Automatic Matching of Terrestrial Scan Data as a Basis for the Generation of Detailed 3D City Models. In O. Altan, editor, *Proc. of the XXth ISPRS Congress*, volume XXXV-B3 of *IAPRS*, 2004. URL: www.isprs.org/istanbul2004/comm3/papers/429.pdf.

[3] M. A. Fischler and R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Comm. of the ACM*, 24(6):381–395, 1981.

[4] L. Grammatikopoulos et al. Automatic Multi-Image Photo-Texturing of 3D Surface Models Obtained with Laser Scanning. In *Vision Techniques Applied to the Rehabilitation of City Centres*, 2004. URL: www.survey.ntua.gr/main/labs/photo/staff/gkarras/Karras_cipa_Lisbon_04.pdf.

[5] S. Heuel. *Uncertain Projective Geometry: Statistical Reasoning for Polyhedral Object Reconstruction*, volume 3008 of *LNCS*. Springer, 2004.

[6] R. Liu and G. Hirzinger. Marker-free Automatic Matching of Range Data. In R. Reulke and U. Knauer, editors, *Panoramic Photogrammetry Workshop*, volume XXXVI-5/W8 of *IAPRS*, 2005. URL: www.informatik.hu-berlin.de/sv/pr/PanoramicPhotogrammetryWorkshop2005/Paper/PanoWS_Berlin2005_Rui.pdf.

[7] D. G. Lowe. *Perceptual Organization and Visual Recognition*. Kluwer, Boston, 1985.

[8] H.-F. Schuster. Segmentation of LIDAR Data Using the Tensor Voting Framework. In O. Altan, editor, *Proc. of the XXth ISPRS Congress*, volume XXXV-B3 of *IAPRS*, 2004. URL: www.isprs.org/istanbul2004/comm3/papers/426.pdf.

[9] M. Wang and Y.-H. Tseng. LIDAR Data Segmentation and Classification Based on Octree Structure. In O. Altan, editor, *Proc. of the XXth ISPRS Congress*, volume XXXV-B3 of *IAPRS*, 2004. URL: www.isprs.org/istanbul2004/comm3/papers/286.pdf.