Joint combination of point cloud and DSM for 3D building reconstruction using airborne laser scanner data.
Fayez Tarsha-Kurdi, Tania Landes, Pierre Grussenmeyer

To cite this version:

HAL Id: halshs-00264859
https://halshs.archives-ouvertes.fr/halshs-00264859
Submitted on 19 May 2008
Joint combination of point cloud and DSM for 3D building reconstruction using airborne laser scanner data

F. Tarsha-Kurdi*, T. Landes, P. Grussenmeyer
Photogrammetry and Geomatics Group, MAP-PAGE UMR 694,
INSA de Strasbourg,
67000 Strasbourg, France. Phone/Fax: +33 3 88 14 47 33
fayez.tarshakurdi|tania.landes|pierre.grussenmeyer@insa-strasbourg.fr

Abstract— More and more cities are looking for service providers able to deliver 3D city models in a short time. Airborne laser scanning techniques make it possible to acquire a three-dimensional point cloud leading almost instantaneously to Digital Surface Models (DSM), but these models are far from a topological 3D model needed by geographers or land surveyors. The aim of this paper is to present the pertinence and advantages of combining simultaneously the point cloud and the normalized DSM (nDSM) in the main steps of a building reconstruction approach. This approach has been implemented in order to exempt any additional data and to automate the process. The proposed workflow firstly extracts the off-terrain mask based on DSM. Then, it combines the point cloud and the DSM for extracting a building mask from the off-terrain. At last, based on the previously extracted building mask, the reconstruction of 3D flat roof models is carried out and analyzed.

I. INTRODUCTION

Digital Terrain Models (DTM) or digital 3D building models as archived in Geographic Information Systems might constitute reference layers for subsequent requests. Nevertheless, a complete automatic processing chain working successfully over the entire areas of interest has not yet been published. That’s why, although they are time consuming, conventional techniques like photogrammetric stereo-restitutions remain in dominant position. This shift occurs at the latest when the model delivered by the service provider has to be checked and corrected.

The suitability of airborne laser scanner techniques for 3D object reconstruction has been proved over the last decade [15]. So, the automatic generation of 3D city models from Lidar data gains importance. To carry out this task, generally two successive steps are considered. First of all, the off-terrain objects are separated from the terrain objects, and then the buildings are extracted from the off-terrain objects.

In this context, the processing based exclusively on point cloud is known to be time consuming. Besides, a processing based on Digital Surface Models (DSM) is damaged by the interpolations made beforehand and represents only a smoothed image of the real geometry.

If the first and last echoes are used together, results are often very satisfying [19], since the second echo helps to distinguish points captured on the top of the canopy from those captured on the ground. But in most cases, the second echo is less accurate than the first one [27] or it is not always available ([24] and [19]). In this context, only the first echo Lidar data is considered.

The topic of this paper is to present a general method for the automatic segmentation and modeling of buildings by the simultaneous combination of the point cloud and the DSM.

In the first step, the city point cloud is segmented automatically into three essential classes: terrain, vegetation and buildings. This operation allows isolating the buildings class which is the input data for the second step. In the last processing step, the building models are automatically constructed based on the building point clouds and on the building mask obtained in the previous step.

The pertinence and advantages of combining both data layers -the raw point cloud and the DSM- are presented through the principal steps of the workflow.

II. STATE-OF-THE-ART

The topics which are regularly envisaged in the research are the automatic point cloud segmentation and the automatic building modeling.

A. Automatic point cloud segmentation

The studies carried out in the field of automatic point cloud segmentation emphasize two main families of treatments: (a) approaches where the support is mainly an image produced by interpolation and/or segmentation of the raw point cloud ([14], [15], [7], [13], [11] [20]). In this case, segmentation mainly the generation of objects composed of similar pixels; (b) approaches trying to concentrate their processing on a point level and where segmentation means the discrimination of several clusters within a point cloud ([18], [10] [17] [21]).

For the segmentation of the data into terrain and off-terrain classes, the algorithm developed in this paper finds its place in the category (a), because a raster DSM and image processing procedures are used.

For the second segmentation chain, the proposed approach joins the category (b) in the sense that it uses the relevant information provided by the original Lidar 3D point cloud.
2007 Urban Remote Sensing Joint Event

B. Automatic building model reconstruction

According to the literature, two kinds of approaches characterize automatic building reconstructions.

The first type of approaches is model-driven, as the approaches suggested by ([14], [25] and [26]). It consists of searching the most appropriate model among basic building models contained in a models library. Main advantage of this concept is the geometric suitability obtained for the buildings under study. In the case of complex roofs, several authors suggest to segment the point cloud into primitives (often by using ground plans) before handling every building independently ([23], [5] and [8]).

The second approach frequently proposed in the literature is data-driven, as the approaches suggested by ([22], [14], [4], [12], [2], [10], [1] and [3]). This method attempts to model a simple or a complex building by using the point cloud as starting data. Thus, after series of more or less complex operations, this technique allows the generation of a model without belonging to a specific library.

The concept of the presented project is based on the second category of approaches. In this way, it constructs in an automatic way a complete 3D description of the buildings. Indeed, at the end of the automatic segmentation, the so-called “normalized DSM” or “nDSM” contains only the buildings. So it combines the point cloud with the nDSM for completely reconstructing the 3D building model. For the moment, the buildings produced through this chain are characterized by plain roofs. The main altitude of the points characterizing the building roof is used as elevation information. In this part of 3D building reconstruction, again the importance and gains of using both point cloud and nDSM are exposed.

III. DATA

In order to evaluate the proposed approach, two point clouds acquired over different areas are available (Table 1).

<table>
<thead>
<tr>
<th>Acquisition</th>
<th>Sensor</th>
<th>Point density</th>
<th>Flight height</th>
<th>Puls. frequency</th>
<th>Field view</th>
<th>Echo Points/Field</th>
<th>Points/ dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>September 2004</td>
<td>TopScan (Optech ALTM 1225)</td>
<td>1.3 points/m²</td>
<td>1440 m</td>
<td>25 kHz</td>
<td>± 26°</td>
<td>450000</td>
<td>400000</td>
</tr>
</tbody>
</table>

Table 1. Characteristics of the point clouds used.

The first test site “Strasbourg centre” is a residential area in Strasbourg main city centre with tangled up houses. The second test site “Victoire boulevard” site is in Strasbourg main city centre with tangled up houses. It contains broad roads and trees near to large buildings.

IV. AUTOMATIC POINT CLOUD SEGMENTATION

The automatic point cloud segmentation consists of isolating the points describing the off-terrain (building, vegetation, trees) and then to extract the buildings from it. The advantages of using both DTM and point cloud are emphasized in the following paragraphs.

A. Use of the DSM for terrain and off-terrain segmentation

A 3D Lidar point cloud is characterized by an irregular distribution in the horizontal and the vertical plan. So to apply digital filters, the dataset has to be transformed into a uniform 2D grid. In order to preserve the real measured altitudes and the validity of the topological original relationships between points, a nearest neighbor interpolation technique is used. Of course, a determinant criterion is the definition of the sampling value (resolution) of the DSM, which has to be related to the mean point density. Thus, one pixel has to contain at least one point. The sampling value SV is deduced from the relation (1).

\[ SV = \frac{1}{\sqrt{\text{density}}} \]  

Once the DSM is calculated, the segmentation procedures begin. The initial concept of segmentation is established on the analysis of height texture measures, i.e. local height detections. So, to detect the off-terrain segment edges, as suggested by [15], directional gradient filters are applied on the DSM with 3x3 kernels under eight different rotations. Thus, eight bands are generated in which the grey values represent height differences. Then, the maximum gradient for each pixel is searched over the eight bands and assigned to a matrix \( AZ \) (equation 2).

\[ \Delta Z_{ij} = \max \,(G_{ij})_k \]  

Where i, j: pixel position in line and column; \((G_{ij})_k\) : kth band of filtered image; k : gradient band number \((k=1,...,8)\).

Figure 1. a) Off-terrain segment edges (matrix \( A \)). b) Off-terrain class. c) Off-terrain after morphological filtering. d) Noise building kernel mask. e) Building kernel mask. f) Building mask.

A first threshold \((S_1)\) is defined according to the smallest detectable building. Thus if \(\Delta Z_{ij} > S_1\), so the pixel \((i, j)\) describes an off-terrain edge and takes the value 1 in a binary matrix called “matrix \( A \)”. At the end of this procedure, the matrix \( A \) presents the off-terrain edges class as shown in Fig. 1a.

At this stage, a convolution operator is applied on matrix \( A \) which has two objectives: the first one consists in filling the body.
of the off-terrain segment borders already detected; the second one consists, in parallel, in assigning a neighboring ground value to the off-terrain pixels that facilitates the generation of a DTM. Finally, the binary matrix \( A \) contains the off-terrain class as illustrated in Fig. 1b and Fig. 1c. Simultaneously, a Test ground matrix is calculated and contains three values: 0 for the terrain pixels, 2 for pixels located inside the body of the buildings and 1 for vegetation and building contour pixels. The details of these operations are explained in [9].

B. Combination of point cloud and nDSM for building extraction

For a reliable automatic detection of the buildings, the joint use of the nDSM and the original point cloud is of fundamental importance in this part of the segmentation. As explained in the following paragraph, their advantages are exploited through the creation of topological criteria.

1) Topological assumptions: The nDSM as well as the original point cloud are considered simultaneously, since the algorithm exploits the fact that one cell or pixel in the nDSM may contain a variable number of cloud points. So we can define for each cell of the nDSM several \( Z \) values Fig. 2.

![Figure 2. Types of Z values occurring in one cell](image)

An example of superimposition of the 3D point cloud on the DSM is presented in Fig 3a. The Fig. 3b shows that the number of points per pixel is variable. This is due to the irregular distribution of Lidar points explained by the relief, the acquisition conditions and the reflectance of vertical elements.

Consequently, three attributes characterizing this local point cloud can be deduced:

- The first attribute presents the real value of the point altitude.
- The second attribute contains the maximum, the minimum and the mean of the altitudes occurring in one cell.
- The third attribute is the DSM altitude value, which presents the raw DSM value obtained by interpolation of the cloud points.

In fact, the real value (first attribute) is unknown in opposition to all other attributes. The second and third attributes can be calculated and allow studying the spatial topological relationships between the points included in a cell. So the maximal vertical distance between the points of the same cell, the number of points per cell Fig. 3e and Fig. 3f and the maximum slope angle can be deduced.

![Figure 3. a) 3D point cloud (with sequence number) projected on the DSM. b) number of points per cell. c) Z values of the cloud points. d) DSM calculated by interpolation. e) Mean of the point altitudes occurring in one cell. f) Maximal vertical distance between the points of the same cell.](image)

If \( \sigma_h \) represents the Lidar data relative accuracy in \( Z \), \( \alpha \) the maximum slope of a plan adjusting a roof locally and \( p \) the DMS sampling value, then the threshold \( S_3 \) regarding the maximum height difference between points in a cell can be expressed in equation (3). Fig. 4 illustrates the maximum slope angle tolerated for the plan fitting a roof.

\[
S_3 = \tan(\alpha) \cdot p + \sqrt{2 \cdot \sigma_h} \quad (3)
\]

![Figure 4. Relationship between \( S_3 \), \( \alpha \) and \( p \).](image)

If more then one point belongs to a cell, the last equation can be calculated.

2) Vegetation elimination using point cloud and DSM: All of these parameters help to describe the pixel nature (vegetation or building) and are consequently useful for the elimination of a large part of vegetation. Indeed, the assumption saying that the roofs are generally composed of plane surfaces helps to improve the distinction between buildings and vegetation.

In order to remove the remaining vegetation segments considered as noise, a threshold based on the smallest foreseeable building surface (\( S_{22} \)) is introduced. In this manner the noise building kernel mask can be cleaned and the remaining segments present now the kernels of buildings (see Fig. 1d and Fig.1e).
3) Building segment completion using the nDSM: The last step of this building extraction task consists in completing the building kernels. This is done with the use of the nDSM on which a specific region growing algorithm working on the 8 neighboring height differences $\Delta h$ is applied. A last filter erases the remaining segments by regarding again the smallest foreseeable building segment (Fig. 1f).

Results obtained through this workflow are very satisfying. Indeed, the major part of the buildings is well classified (Fig. 1f). Only a few pixels of vegetation are misclassified and easily rejected by mathematical morphology operators. The explanation for the remaining misclassification is that a group of points, within the vegetation class, accepts an average plan with negligible residues and respects the whole topological assumptions.

C. Thresholds used as segmentation parameters

In the presented segmentation, three thresholds are used:

- $S_1$ represents the minimum height difference between terrain and off-terrain which is defined according to the lowest building elevation.
- $S_2$ represents the smallest foreseeable building surface.
- $S_3$ represents the maximum height differences $\Delta h$ in the neighborhood, on a roof.

The $S_1$ threshold is used to separate the terrain from the off-terrain class. Its value is correlated to the sampling value of the DSM. So when the mean distance between two neighboring points increases, the threshold $S_1$ has to increase too. Generally, $S_1$ value is set between 4 and 6 meters.

The $S_2$ threshold is used three times in the segmentation algorithm. Firstly, it is inserted during the terrain/off-terrain segmentation, in order to eliminate the noise and the small segments from the off-terrain class. Secondly, it is used to clean the building kernel mask from the remaining vegetation segments in the off-terrain class. Finally, it is introduced to verify the absence of vegetation segments in the building mask obtained at the end of the processing. Its value is related to the urban typology nature, so it lies between 50 and 150 m².

The $S_3$ threshold is used in the region growing algorithm during the segmentation step in which vegetation and buildings are separated. It is used to test the eight neighboring pixels of each pixel in order to detect if it belongs or not to a building kernel. In general, the value of $S_3$ has to be smaller than the value of $S_1$ since we are looking for smaller height differences. The value of $S_3$ can be estimated using the maximum foreseeable roof slope as explained in Fig. 4 and it takes values between 0.3 and 1.6m.

At now, buildings are extracted from the off-terrain class and need to be modelized by geometric reconstruction.

V. AUTOMATIC BUILDING RECONSTRUCTION

The input data for the 3D building modeling are the 3D point cloud and the building mask resulting from the automatic segmentation explained previously.

A. Contour polygon extraction using the nDSM

Using a region growing algorithm, it is possible to isolate automatically every building in a binary normalized surface model (nDSMb). Simultaneously, the points belonging to that building are extracted from the general point cloud. Thus, every building can enter into the process individually.

In the first step, the algorithm extracts the polygon describing the building contour from each nDSMb. In each nDSMb, a pixel of the building contour is equal to 1 and at least one of its neighbors has the value 0. This remark enables to detect easily the pixels belonging to a building contour.

Fig. 5a presents the case of a building characterized by an inner court. In order to extract both inner and outer contours, a 3 by 3 moving window (Fig 5b) runs over the nDSMb to obtain the result shown in Fig. 5c. Then, the interior and exterior polygons have to be separated by a region growing algorithm (Fig. 5d and Fig. 5e).

Figure 5. a) nDSMb of a building defined by two contour polygons. b) Moving window used to extract the contour polygons. c) Extraction of the building contour polygons. d) Building contour polygon number 1. e) Building contour polygon number 2.

The result of this extraction is a raster polygon with non rectilinear sides and with parasitic pixels (Fig. 6a).

B. Rectification of the contour using the nDSM

At this stage, it is necessary to exclude the non representative points and to attenuate the noise along the buildings contours. The condition that each pixel of the polygon must have only two neighboring pixels allow to clean the polygon sides as illustrated in Fig. 6b. Nevertheless, the rectilinear sides are not yet obtained.

Once the building contours are improved, the contour polygon of the building is decomposed according to its facades. To achieve this goal, the Douglas Peucker technique is used as proposed in [6], [25] and [1]. The application of this technique provides the coordinates of the extremity points as well as the list of all points describing every facade (Fig. 6c).

In the next step, the 3D building facades must be fitted by straight lines in order to reach a realistic geometry in the horizontal plain. For this purpose, the least squares theory is applied to calculate the mean line of each segment.
However, it occurs sometimes that several point lists are obtained for the same facade. This is due to the presence of noise points which are far from the calculated adjustment plan. The decision to merge two contiguous or neighboring segments is taken if one or both of the following situations occur:

- If the acute angle between two contiguous segments is lower than a certain threshold (for example 25°), then the mean line fitting the points composing the two neighboring segments is calculated.
- If the standard deviations are lower or equal to a certain threshold (for example 1 pixel), then the two segments represent only one line, and they are merged.

Otherwise, the segments represent two independent facades. Finally, the most probable facade plains are computed respecting the least squares fitting theory. Fig. 6d presents the segments and the resulting fitting plans.

C. Construction of the 3D building model

The coordinates of the building corners are calculated using the intersection between every two adjacent facades (Fig.6e). Finally the 3D building model is deduced by supposing that the whole building facades are verticals (Fig. 6f).

The buildings produced through this chain are characterized by plane roofs. The mean building height is simply set to the average height of the building cloud points.

No geometric constraints are used to respect the parallelism or the perpendicularity conditions in the final model. Indeed, the use of this kind of geometric constraint generates considerable errors despite some improvements. Nevertheless, the decision to add or not constraints will be let to the appreciation of the user and can be introduced at the end of the process.

VI. RESULTS AND DISCUSSION

The complete workflow i.e. the segmentation as well as the building reconstruction has been applied on several test sites, with several Lidar data characterized by different densities and for areas of different urban typologies (see paragraph II). Fig. 7c presents the buildings extracted for the “Victoire boulevard” test site; and Fig. 7d shows the final 3D model produced.

In order to evaluate the precision of the building extraction and the 3D building model, an estimation method summarized by [16] has been applied. According to [16], the main idea of this method is to transform the evaluation of building models into the evaluation of a classification by discretising the space in pixel (2D).

This method proposes to calculate many factors which describe the building model accuracy in comparison with a reference model. The reference model used for evaluating our results is a topographic plan.

The quality measures used here are the following factors: the “quality rate ($\rho_q$)”, the “branch factor ($\rho_b$)”, the “miss factor ($\rho_m$)”, the “false alarm rate ($\rho_f$)”, “type 2 error ($\beta$)” and the “shape dissimilarity ($\rho_s$)”. Table 2 presents the mean values and the standard deviations obtained for every factor.

Satisfying results mean that the two first factors are near to 1, whereas the five last factors reach the value 0 [16]. The values obtained in Table 2 are explained by the weak point cloud
density, the generalization used in the Douglas Peucker technique, the point cloud homogeneity and at last the buildings lost during the automatic segmentation [9].

Another important reason is the incoherence between the data used for the model (Lidar point cloud data) and the data used for the reference model (conventional topographic maps or photogrammetric restitutions). Indeed, conventional topographic measurements are taken at the bottom of the buildings, whereas Lidar systems capture points on the roofs.

<table>
<thead>
<tr>
<th>ρ d</th>
<th>ρ q</th>
<th>ρ b</th>
<th>ρ m</th>
<th>ρ t</th>
<th>f</th>
<th>ρ s</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.03</td>
<td>0.01</td>
<td>0.1</td>
<td>0.17</td>
<td>0.09</td>
<td>0.12</td>
<td>0.21</td>
</tr>
<tr>
<td>S.D.</td>
<td>0.12</td>
<td>0.11</td>
<td>0.09</td>
<td>0.25</td>
<td>0.08</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Table 2. Results of the precision evaluation factors.

Moreover, in the case of the photogrammetric processing, the details of the building facade are not restituted (like balconies, terraces), whereas these details are contained in the Lidar point cloud and are taken into account in the generated model. If the building facade does not contain significant details, the adjusted plan is calculated correctly, in coincidence with the reference model.

Generally, the proposed approach provides very good results, for both point densities used. Nevertheless, it shows also some limitations. Indeed, the segmentation result is less reliable when trees and buildings are close to each other and are of the same height. Furthermore, if a block of buildings is composed of several houses with similar heights (falling under the threshold \( S_1 \)) the algorithm will misclassify the off-ground pixels as terrain instead of building.

Concerning the building reconstruction, a way to improve it is to introduce shape or geometric criterions. Moreover, the modeling algorithm shows limitation in the case of building with several connected contour polygons. This occurs for example when the point cloud is heterogeneous. The solution consists in recalculating the nDSM from the original building point cloud using a higher sampling value.

VII. CONCLUSION

The most relevant idea in the approach presented here is to combine the original point cloud and the generated DSM at strategic moments of the automatic segmentation and building reconstruction. In this context, the Lidar data are considered at a point level, pixel level, segment level and global level during the processing chain. However, although the process returns regularly to the raw point cloud, the computing time required for extracting buildings is almost negligible.

Of course, this methodology may still be improved to particular cases, since some small vegetation segments may remain as well as a few buildings can be lost until the end of the process. For the automatic building reconstruction, the building mask and point cloud are also used jointly.

The results obtained by this combination are very promising since the accuracy analysis validates the processing chain for a majority of buildings. At this stage, the reconstruction of the building roofs constitutes the forthcoming modeling phase.

REFERENCES


[12] G. Vossel
sman, S. Dijkm


[19] L. Wotruba, F. Mor


