Acquisition of 3D regular prismatic models in urban environments from DSM and orthoimages

João Ferreira & Alexandre Bernardino
Instituto Superior Técnico, ISR Lisboa, Portugal

Keywords: Urban modelling, regular polygons, model selection, 3D reconstruction, digital surface maps (DSM), orthoimages.

ABSTRACT: We present a novel method for fitting polygons to segmented image regions. The method is applied to the automatic reconstruction of 3D prismatic building models in urban areas, using regions segmented from the Digital Surface Map (DSM) and color orthoimages obtained from Google™-Earth. Polygon selection methods with complexity penalty terms are developed for the acquisition of simplified building outline models. Regular polygons such as rectangles, right trapezoids and general trapezoids are most common in urban areas, and special methods are developed for fitting these shapes to the segmented regions. Notwithstanding, irregular polygons are also considered since some ancient urban areas are plenty of buildings with non-standard shapes. Results are presented on the 3D reconstruction of the city of Lisboa, Portugal, with simple prismatic models.

1 INTRODUCTION

In this paper we address the problem of automatic generation of 3D building models in urban areas. Several cities around the world have built virtual models of certain parts of their territory. These virtual models are important for many purposes, e.g urban modelling, telecommunications for transmitter placement, environmental planning, simulation etc. However, the large majority of the models are obtained by manual methods, requiring many man-hours of work for introducing data which makes the creation and maintenance of such models too costly and time consuming. Therefore there is a great deal of effort put in the research for automatic or semi-automatic tools for the acquisition, validation and update of cadastral data in urban areas. Classically the main input data for the production of 3D city models were aerial images, terrestrial images, map data, and surveying data. In the last years LIDAR (LIght Detection And Ranging) has become a very attractive alternative for the acquisition of 3D information since this technique directly provides a high density of 3D points. However, the obtained maps are usually smoothed and irregular, and should be complemented with other sources of information. Several works have proposed the use of aerial images to complement the LIDAR maps and present methods to combine their information toward 3D building reconstruction [3,4]. Our work fits in the same approach but the main problem addressed here is the acquisition of regular shapes for buildings' outlines. This is important in order to obtain succinct representations of buildings, allowing light storage cost and fast manipulation in visualization systems. Figure 1 shows a diagram with the overall processing pipeline for our reconstruction method.

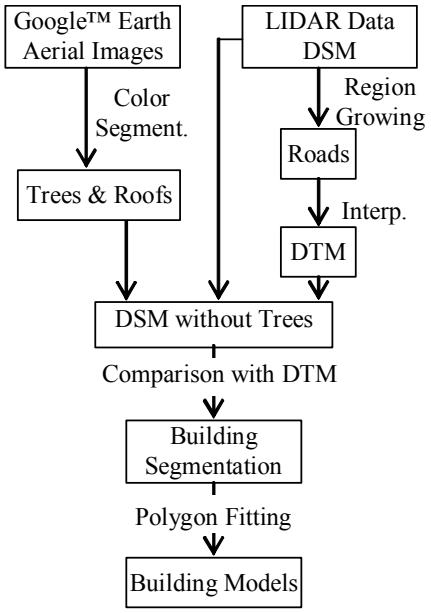

Figure 1. Architecture of the proposed building reconstruction method. In future work the information of building roofs will be used both for texture mapping and roof modeling and reconstruction.

We use LIDAR data in the form of a digital surface map (DSM), and color orthoimages obtained from Google™-Earth. In section 2 we describe how these

two sources of information can be effectively used for the segmentation of buildings' regions, including the between-map registration problem, DSM normalization, removal of vegetation areas, and map binarization. A first step consists in identifying roads in the DSM by region growing processes. The Digital Terrain Map (DTM) is then obtained by interpolating a smooth surface supported in the identified roads. This allows a first segmentation of the buildings by comparison to the DSM. However, spurious segmentations are obtained due to vegetation.

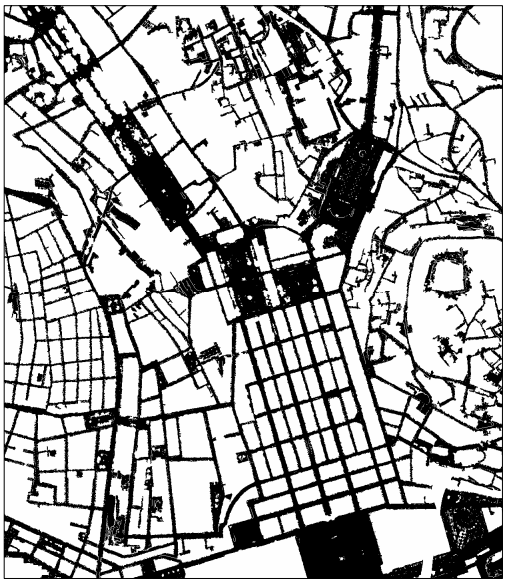

Figure 2 – Result of region growing on the DSM. Dark areas correspond to the roads and ground plane regions. Bright areas correspond to buildings, trees and other above ground structures.

A color segmentation processing on the orthoimages allows the detection of such cases and removing most of the segmentation errors. After this segmentation process we obtain a binary map with regions corresponding to buildings but these regions are still very irregular to be used for 3D reconstruction - they would render very unrealistic building models, as well as requiring a large amounts of information to represent them. Therefore, in section 3 we present methods to obtain simplified representation of the buildings' outlines consisting regular polygons whenever adequate. This is a reasonable assumption since most buildings in urban areas follow simple geometric models. In [1], buildings are modeled by rectangles or combinations of rectangles. Instead we define a larger set of 2D regular geometric shape hypothesis with different complexities (from rectangular to trapezoidal shapes). These models are fit in the segmented data, using as criteria, both a goodness of fit term and a complexity term. Initial shapes for the regions are obtained with some heuristics and a final iterative minimization procedure is used to select the best hypothesis. The introduction of a complexity penalty term is important to reduce the amount of information required to represent the models, thus allowing faster visualization and lighter hardware requirements. If none of the predefined shapes fit well in the regions, the classical Douglas-Peucker line simplification algorithm [2] is used to fit a general polygon to the building's boundary. In section 4 we explain the model selection procedure, i.e. the choice of the best polygon, regular or non-regular, to represent a certain building. The obtained polygons are then used to create 3D virtual prismatic models of the buildings by extruding the 2D shape to a height provided by DSM information. Along the paper we present the results obtained by applying the different phases of the processing pipeline to available DSM data from Lisbon city (Portugal) and orthoimages obtained from Google™-Earth. In section 6 we present the final results obtained illustrating the performance of the proposed polygon fitting and selection mechanisms. Finally, in section 7 we present the conclusion of this paper and ideas for future work.

2 INITIAL SEGMENTATION

To perform the acquisition of regular prismatic models for blocks of buildings[1], the data provided by DSM and aerial orthoimages is processed to create and initial binary image indicating the regions belonging to buildings with high likelihood. In the next lines we succinctly describe each of the modules of the processing pipeline shown in Figure 1.

2.1 *Obtaining the Digital Terrain Model*

Digital Surface Model data is used to obtain the DTM. First, a region growing algorithm, based on height differences, is applied in the DSM, thus getting a representation of all connected streets and open spaces from initial seed points used in the region growing algorithm. The region growing algorithm is applied several times for different initial seed points and, by joining the several results, a good representation of all the streets and open spaces is obtained. Results of the region growing procedure applied to a part of the Lisbon city DSM are shown in Figure 2. After the ground level points have been determined, the DTM is obtained by interpolating a smooth surface supported on the ground level points (roads and open spaces with its heights obtained from the DSM).

2.2 *Color segmentation*

The color segmentation of the orthoimages is based on the Mahalanobis distance of a color pixel to a single Gaussian color model [5].

[1] In this paper we do not consider single city buildings but connected blocks of buildings. Whenever we apply the word "building" it should be understood a "block of buildings".

For the roof detection, the RGB color model is the one with the most Gaussian profile (for both light and dark roofs) while for the vegetation we use the UVL color model.

After the segmentation process, two masks are obtained, one for the roofs and other for the vegetation. Note that the roof mask is only used for future work. Results of the vegetation segmentation are presented in Figure 3.

Figure 3. Green color segmentation. Dark areas correspond to regions containing vegetation with high likelihood.

2.3 *Building Segmentation*

Comparing the DSM and DTM, a first segmentation is obtained based on the heights of structures, which include spurious segmentations due to vegetation.
With the application of the vegetation mask on the first building segmentation, followed by some morphological operations and elimination of small area objects, a good building segmentation is then obtained. Results of this process are shown in Figure 4.

3 SHAPE MODELING

For each one of the objects in the final building segmentation, four types of shapes are considered: regular polygons like rectangles, rectangular trapezoids, trapezoids and general non-regular polygons.

The first step for the modeling process is to obtain a contour representation of the object which is made by edge detection using the canny method, followed by the arrangement of the boundary pixels. Then, polygons of several types are fit in the contour, each one with a different number of edges. The end result of this procedure is to associate to each building a set of polygons with a number of sides ranging from 4 to a number determined by the Douglas-Peucker algorithm [2]. Then, in the shape selection step (section 4), we will choose the best model taking into account the fitness error and a complexity term proportional to the number of edges.

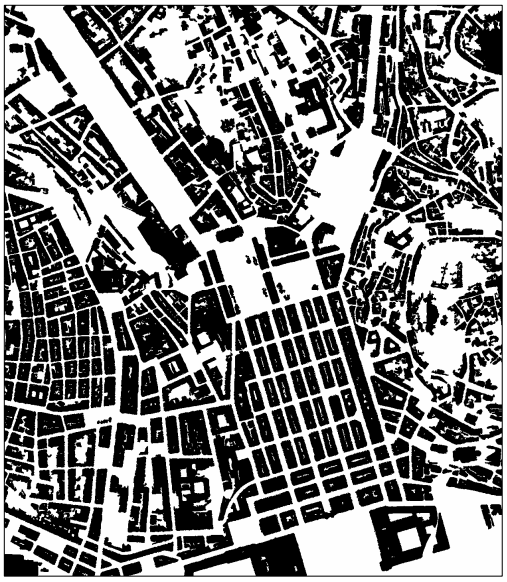

Figure 4 – Building Segmentation. Dark areas correspond to blocks of buildings, with high likelihood.

In both the shape modeling and the shape selection phases, the overlap error is used either for shape fitting and selection. Let us define the *overlap error* as the sum of pixels inside the estimated polygon but not on the object, plus the sum of pixels in the object but not inside the contour, is then associated to each one of the contours estimated.

3.1 *Non-regular polygon estimation*

The non-regular contour is modeled by a modified Douglas-Peucker algorithm, which also returns the cost associated to each point and is depth [2]. This extra information is then used for iteratively removing one edge at a time of the polygon: the point with the lower cost is removed at each iteration. This allows to obtain simpler models of the general contour (Figure 5), ranging from the original, non-simplified contour, to a quadrilateral (four corner polygon). All polygons estimated in this sequence are stored and the *overlap error* is associated to each one for posterior use.

3.2 *Regular polygon estimation*

In this work we consider that most of the buildings' outlines have regular shapes and have developed specialized algorithms for fitting rectangles, right trapezoids and general trapezoids to the segmented contours.
The fitting algorithms consist in an initialization phase based in the region first and second order moments (centroid, orientation, major and minor

axes length), and an iterative phase where the length of the polygon edges and its orientation are refined in an interleaved way.

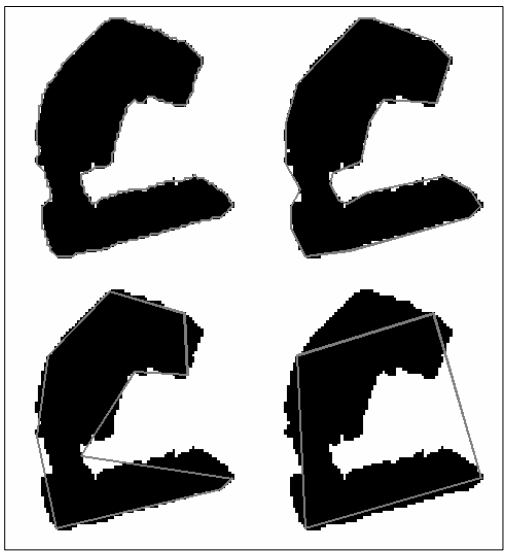

Figure 5. From right top to left bottom: initial contour; estimated contour using Douglas-Peucker algorithm; contour simplified to 10 corners; contour simplified to 4 corners.

For rectangles, an initial estimation is made based on the image blob's first and second order moments (centroid, orientation, major and minor axes length), which roughly indicate the its position, orientation and dimensions. For the trapezoids, the initial estimation is made based on the first four corners returned from the Douglas-Peucker algorithm.

In the iterative phase, the position of the edges and the overall orientation of the polygon are updated step by step. The coordinates are first normalized so that two parallel sides are horizontal. Then, for each side of the polygon, it is associated one error measures at each edge (or half edge for the non-orthogonal sides in the trapezoids). The error measure is the average difference between the edge and the closest point in the blob orthogonal to the edge. Blob points to the inside of the current polygon have a negative contribution to the error measure while points to the outside have a positive contribution. Then each edge vertex is shifted orthogonally to the edge orientation an amount proportional to its error measure. If the error is positive the vertices are moved to the outside of the polygon and *vice-versa*. The trapezoid non-orthogonal sides are splint in two halves and each vertex has its own error measure, being is shifted independently of the other. Let us analyze more closely the more complex case of a general trapezoid, as shown in Figure 6. Six error measurements are obtained: up edge (E_u), down edge (E_d), lower part of left edge (E_{ld}), upper part of left edge (E_{lu}), lower part of right edge (E_{ru}) and lower part of right edge (E_{rd}). Then, the horizontal (x) and vertical (y) coordinates of the trapezoid vertices are updated according to the following expressions:

$$\begin{cases} V_{lu}^x = V_{lu}^x + E_{lu}\,, & V_{lu}^y = V_{lu}^y + E_u \\ V_{ru}^x = V_{ru}^x + E_{ru}\,, & V_{ru}^y = V_{ru}^y + E_u \\ V_{ld}^x = V_{ld}^x + E_{ld}\,, & V_{ld}^y = V_{ld}^y + E_d \\ V_{rd}^x = V_{rd}^x + E_{rd}\,, & V_{rd}^y = V_{rd}^y + E_d \end{cases}$$

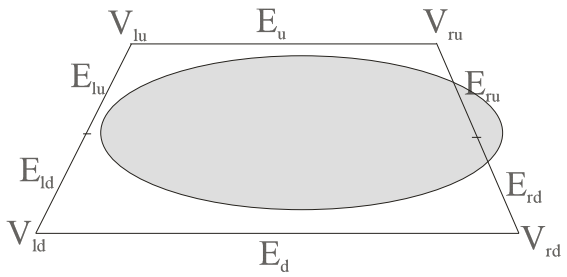

Figure 6. Error measurements computed in a general trapezoid.

After the edges have been updated, the overall orientation of the polygon is refined by rotating the estimated contour left and right and evaluating the evolution of the *overlap error* to decide whether to update the orientation at this step or not. Examples of intermediate steps and final estimation are shown on Figure 7.

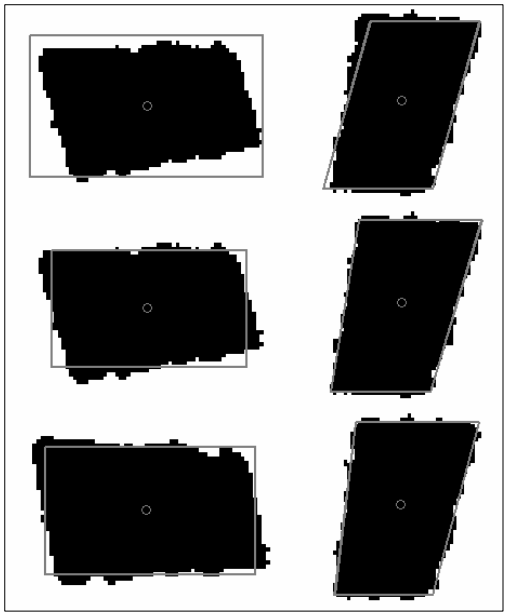

Figure 7. First column – rectangle; Second column – trapezoid. Rows top to bottom: initial estimation; edge size update; orientation update.

4 SHAPE SELECTION

After the four different contours are estimated for each one of the objects, a selection is made, based on

the overlap error and its complexity so the best model is chosen to represent each object.

For the non-regular polygons one must first determine the best compromise between overlap error and complexity for each one of the different estimated models. First the overlap error is normalized by the area of the object and then cost function is computed based on the normalized overlap error and complexity of the model (number of edges). The best non-regular polygon is the one with the minimum cost.

For regular polygons, the cost function is computed similarly where complexity is not directly associated with the number of edges but with the number of degrees of freedom (rectangle-4, right trapezoid-5, general trapezoid-6). However the proportionality factor affecting the complexity term in the cost function of regular models is lower than for non-regular models in order to favor the selection of regular polygons. This parameter is manually tuned based on visual inspection of the results.

The cost functions of all model hypotheses for a given building are compared and the one with the minimum value is the model chosen to represent the object. Manual corrections are possible using a simple graphical user interface (GUI).

5 RESULTS

The whole processing pipeline is applied to available DSM data from Lisbon city (Portugal) and orthoimages obtained from Google™-Earth. Results are presented in Figure 8 illustrating the output of the algorithm.

In Figure 4, after the application of the vegetation mask sometimes important information is removed because some trees are hiding important structures. This is not a simple problem to be solved automatically and will be included in future work. Also some trees aren't well removed in the tree removal process mainly because the DSM data is older than the orthoimages.

The automation of the process depends on some user defined constants: threshold on region growing, DSM/DTM comparison threshold, color segmentation threshold, control constants in the estimation algorithm and finally the complexity weight constants. While the first tree constants have great impact in the detection of buildings, the last ones have their impact in the classification of buildings.

In general, buildings are well estimated and classified. The only exceptions are some buildings with small non rectangular corners, small buildings and trapezoidal buildings with near rectangularity which are classified as rectangular (the last ones can be manually adjusted).

Applying the fitted polygons' mask of Figure 8 in the original orthoimage, the number of false positives (where ground or trees are inside the building area) is very low aside the patios inside the buildings. In the other hand, applying the inverted mask, the number of true negatives (where buildings are not classified as buildings) is higher, although some of them are due to the fact that when the DSM was obtained the buildings still didn't exist. The main sources of true negatives are the tree removal process and the DSM/DTM comparison.

Figure 8. Fitted polygons to all building regions. Rectangular buildings are shown in black, trapezoidal buildings in dark grey and general polygons in light grey.

In general, the computation time per building estimation is relative quick although larger objects take more time to estimate.

Finally, the 3D reconstruction is made by extruding the 2D model to a height equal to the mean of the building height without patios, which is provided by the DSM and the building segmentation, as shown in Figure 9.

6 CONCLUSIONS AND FUTURE WORK

In this paper we have presented a semi-automated method for the acquisition of simplified 3D models of urban areas from DSM and aerial image data. The whole processing pipeline is presented and illustrated with intermediate results of the several processing steps and applied to the 3D reconstruction of a significant part of the Lisbon city. We present the methods used for road detection, building segmentation, vegetation detection and 3D reconstruction, but special emphasis has been put on the modeling of the buildings with regular prismatic models, allowing succinct representations of the data and reducing the complexity of managing its storage and visualization. The overall process has good results and the level of automation is very high (just needs some su-

pervision of the results and adjustments of the type of polygon).

In future work several issues should be addressed. Since many buildings can be represented by the composition of regular shapes, we will study the possibility of such models. For instance an L shaped building could me modeled by two connected rectangles.

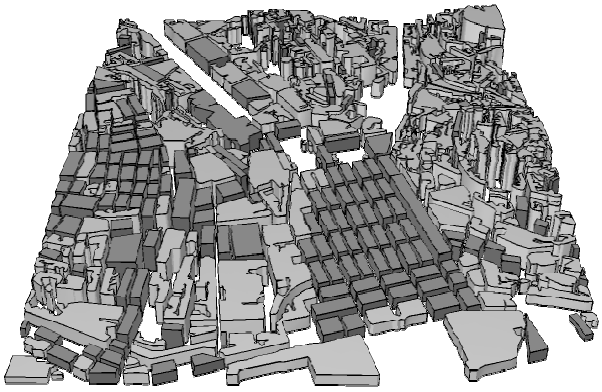

Figure 9. Reconstructed 3D model of Lisboa downtown with simple prismatic models.

Also many buildings on the city have inner patios and the inner building contour should also be considered. Finally, since the results presented in this work reconstruct blocks of buildings instead of individual buildings, in future work we will evaluate the possibility of segmenting blocks into buildings looking for the characteristics of roofs both from aerial images, inspecting the edges, and from DSM data, adjusting planes to the roofs.

7 ACKNOWLEDGMENTS

This work was partially funded by the Portuguese Foundation for Science and Technology (FCT) through project POCTI/AUR/48123/2002 – Spatial validation of complex urban grids in virtual immersive environments.

8 REFERENCES

[1] S. Vinson, L. Cohen and F. Perlant. Extraction of Rectangular Buildings in Aerial Images. In Proc. Scandinavian Conference on Image Analysis (SCIA'01), June 2001, Bergen, Norway

[2] D.H. Douglas and T.K. Peucker. Algorithms for the reduction of the number of points-required to represent a line or its caricature, The Canadian Cartographer, 10, (2), 112–122(1973).

[3] Martin Huber, Wolfgang Schickler, Stefan Hinz, Albert Baumgartner. Fusion of LIDAR Data and Aerial Imagery for Automatic Reconstruction of Building Surfaces. 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, 2003.

[4] F. Rottensteiner, J. Trinder, S. Clode, K. Kubik and B. Lovell. Building Detection by Dempster-Shafer Fusion of LIDAR Data and Multispectral Aerial Imagery. In Proc. ICPR'04

[5] Jean-Christophe Terrillon, Mahdad N. Shirazi, Hideo Fukamachi and Shigeru Akamatsu. Comparative Performance of Different Skin Chrominance Models and Chrominance Spaces for the Automatic Detection of Human Faces in Color Images. Proceedings of 4th IEEE International Conference on Automatic Face and Gesture Recognition, pp. 54-61, 2000.