Multiple Rectangle Model for Buildings Segmentation and 3D Scene Reconstruction

Samuel Vinson[1] Laurent D. Cohen[2]

[1]EADS Matra Systèmes & Information, now with IDP, iTV company, Paris, France
[2]CEREMADE Université Paris 9 Dauphine, 75775 Paris cedex 16, France
E-mail: vinson@idp.fr, cohen@ceremade.dauphine.fr

Abstract

This paper introduces a method for automatic extraction of buildings in aerial images. We first present a method based on rectangular buildings, which are the most common constructions. After a rough segmentation, we estimate a criterion of similarity of each region with the best matching rectangle. For buildings of complex shapes, we introduce an iterative way to divide a region in order to optimize its approximation by a set of rectangles. We use a parametric deformable model for refining rectangle size and positions. These rectangles are then used to enhance a 3D realistic reconstruction of the scene including building models.

1 Introduction

Realistic models for 3D reconstruction of a scene are increasingly needed in the civil as well as military fields (virtual reality, telephony infrastructure, impact studies, video games,...). From aerial images, Digital Elevation Models (DEM) and then ortho-image, the vertical view of the scene, are computed (Fig. 1-left).

Automatic building modeling has proven to be a difficult task. There has been active research on this subject these last years [2]. Several authors intended to improve building rendering in the DEM. Vestri [6] improves accuracy of DEM by modifying the method of generation by correlation especially on the building frontages. Lee et al. [5] present a semi-automatic system to generate 3D models with rectilinear hypotheses. Kim et al. [4] make use of multiple images to obtain complex models.

In our approach, we want to minimize the operator workload by making completely automatic the 3D modeling of most buildings. Our approach has advantage above previous ones to be less dependent on initial segmentation of above-ground structures. Above-ground structure extraction, like vegetation and constructions, (see Fig. 1-middle) is carried out on the DEM by the algorithms presented in [7].

2 Rectangular Buildings Segmentation

Our first goal is to segment automatically all rectangle buildings from the aerial image. For each blob of Fig. 1-middle, we first find the best matching rectangle and then compute a criteria to test whether the blob is indeed a rectangle.

Rectangle Parameters. We intend to model each above-ground element by a rectangle. A rectangle is completely defined by its center of mass (X_g, Y_g), orientation θ, length L and width l. The rectangle center of mass is the same as the above-ground blob. Its orientation is defined by the principal axes computed over the blob itself by inertia moments. We make an estimate of the size of the rectangle by assuming the blob is a rectangle. Based on the orientation of the principal axis, the eigenvalues λ_+ and λ_- and the second order moments, we find the parameters of the best fitting rectangle. We showed that for a given blob of eigenvalues λ_+, λ_-, the sizes of the best rectangle are obtained by $L = \sqrt{12\lambda_+ + 1}$ and $l = \sqrt{12\lambda_- + 1}$. In the case of a square, the principal axis cannot be computed. However, the orientation of a square shape can be computed by using Fourier descriptors [7].

As seen in middle of Fig. 1 and Fig. 2, some above-ground blobs have not a shape similar to a rectangle.

Rectangle Criterion. Our criterion of similarity is based on a comparison between sets. We selected the Hausdorff measure $\mathcal{H}$ among the criteria we studied [7]. It is equal to the ratio of the intersection area of the two sets $\mathcal{A}, \mathcal{B}$ to the area of their union: $\mathcal{H}(\mathcal{A}, \mathcal{B}) = \frac{\#\mathcal{A}\cap\mathcal{B}}{\#\mathcal{A}\cup\mathcal{B}}$ where $\#X$ is the area or number of elements in the set X. When two sets are equal, $\mathcal{H}(\mathcal{A}, \mathcal{B}) =1$. On the contrary, as two sets tend to differ, their intersection decreases whereas their union increases, resulting in a Hausdorff measure close to 0.

We illustrate in Fig. 1-right the selection by Hausdorff measure for blobs whose size is over 300 pixels. We note that the selection is correct: all blobs that are not rectangular are rejected and only some blobs that could be estimated by a rectangle are excluded.

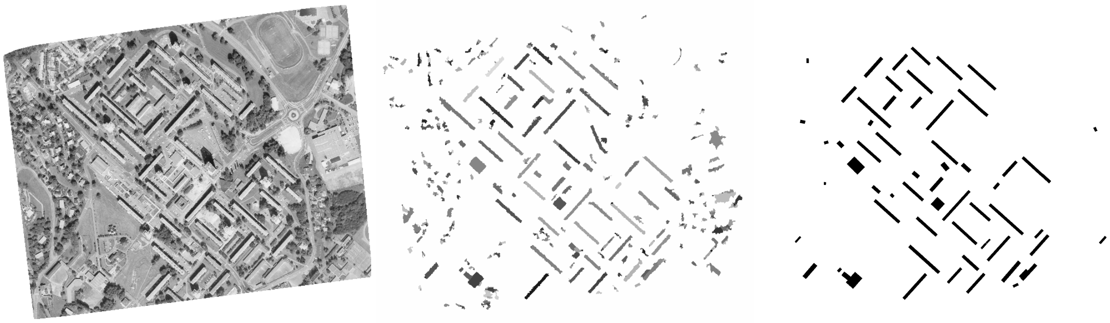

Figure 1. Segmentation of Rectangular Buildings. Left, the ortho-image, middle, the above-ground areas, and their rectangle approximation on the right, when criteria of similarity is satisfied.

Figure 2. Examples of complex buildings and their above-ground blobs.

3 Segmentation of Complex Buildings

A rectangular model for buildings is not always sufficient, as shown in Fig. 2. Therefore we have introduced a new method that enables to divide a shape into a set of rectangular shapes. We want to minimize the number of rectangles and the overlap between rectangles and to maximize the size of rectangles. Therefore we split iteratively a blob in two regions only and find the best way to split in order to get at least one of the two rectangles that gives a very good estimate.

The idea comes from the fact that assuming the blob is a combination of two rectangles, if we cut the shape through the center of inertia in the direction orthogonal to its longer inertia axis, it is likely that one of the two shapes thus obtained is a rectangle. Therefore, as illustrated in Fig. 3, we propose a first split obtained by the line of inertia for the whole blob as defined for the best matching rectangle. We denote by B the complete blob and B_1 and B_2 the two parts of the blob as split by the chosen line. For each blob, the best rectangles (noted R, R_1 and R_2) are obtained as in previous section. The axis chosen corresponds to the smaller eigenvalue. This means cutting the shape across the longer eigenvalue axis. In fact the choice of the axis is based on relative orientation of R, R_1 and R_2.

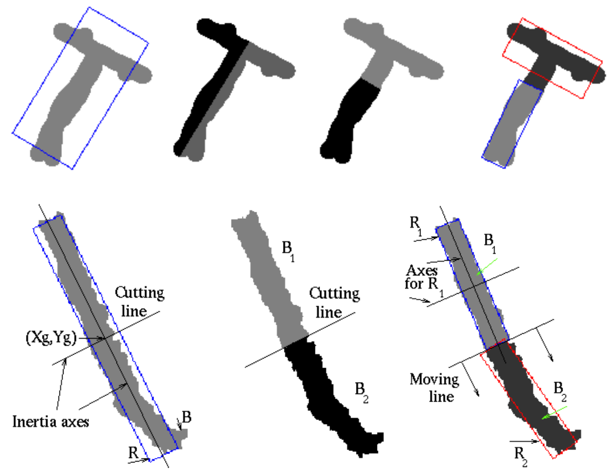

Figure 3. Complex Buildings: First step of splitting the above-ground blob. Above from left to right: best rectangle for a synthetic T-shape blob, splitting the blob though the two axes of inertia, best rectangle for the two regions obtained initially. Below, the same for the complex building shape of figure 2-left.

A second step consists in sliding the splitting axis along the orthogonal line in order to find the best place to cut the blob. As assumed above, one of the two regions should be similar to a rectangle, but we would like to get this matching rectangle as long as possible. A global Hausdorff measure is computed between the union of the two matching rectangles and the complete blob : $\mathcal{H}(R_1 \cup R_2, B)$. The goal is to have the highest global measure, and therefore to optimize the splitting process. Specifically, we proceed by translating the orthogonal line in the direction which increases the area of the region with best rectangular estimation. This

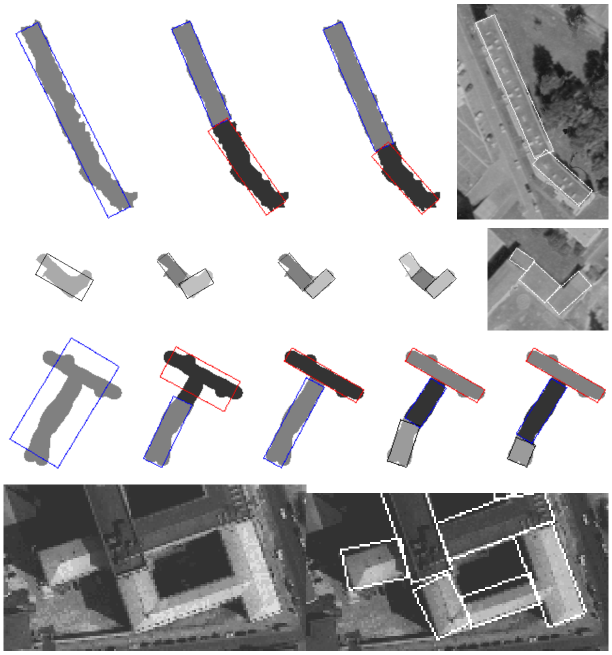

Figure 4. Complex Buildings: Intermediary steps for the splitting. For each example, from left to right the blob and its best rectangle match, the initial split in two regions, and evolution of this splitting or further splitting till reaching equilibrium on the right. Another example on real data is shown on last line.

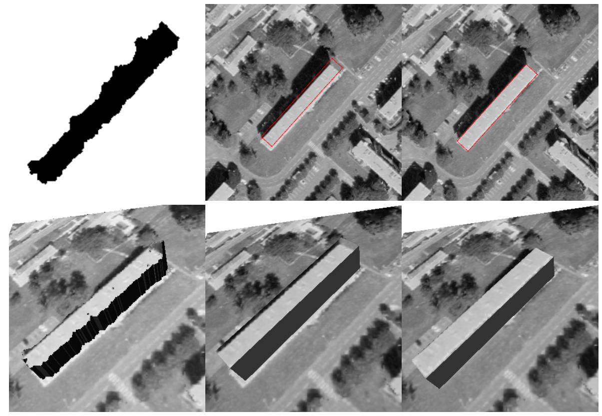

Figure 5. Above: initial blob from DEM, initial rectangle and final result after minimization, superimposed on the ortho-image. Below (see text), compare the 3D rendering using the refined rectangle model to raw blob.

means the rectangle for which $\mathcal{H}(R_i, B_i)$ is larger, say R_1. As shown in Fig. 3, the direction of the cutting line is changed to be the axis of rectangle R_1, which is much more precise than the axis of R. Once the best value of the criteria is reached, we freeze the rectangle approximation R_1 for the good blob $\mathcal{B}_1$.

If the global Hausdorff measure is not high enough, the splitting is repeated on the badly estimated blob B_2. This means we repeat the two steps starting by splitting B_2 by a line and approach each sub-blob by a rectangle. Further details are given in [7].

Figure 4 shows the evolution of the process involved for three examples and the final result for decomposition in rectangles. This shows how estimation results are improved by using this approach for buildings of complex shapes.

4 Deformable Rectangle model

Previous computations result in a rectangle or a set of rectangles per above-ground region. When a rectangle is projected onto the ortho-image, its borders may not fully fit the boundary of the building. We use a deformable rectangle template, as presented for example in [9]. In our case the template is a rectangle defined by its five parameters: coordinates of center X_g, Y_g, orientation θ and sizes L and ℓ. These parameters define the parametric model that evolves by minimizing an energy defined on the four sides of the rectangle $E = E_{01} + E_{12} + E_{23} + E_{30}$, where the vertices (X_i, Y_i) are indexed from 0 to 3. Each term of energy is the integral of a potential P over a side. Energy is minimized by means of gradient descent (see [7] for details). The potential function P is defined as the opposite of the gradient norm of the ortho-image [3]. We then apply Gradient Vector Flow [8] in order to extend the borders power of attraction (Fig. 5-above).

5 3D Reconstruction

One of the goals of this work is to obtain a precise 3D reconstruction of the scene including our models for the buildings. Thus once we got a precise estimate for buildings, we separate the surface reconstruction of the ground regions and above-ground regions. For each rectangle of a building, we use a parallelepiped model. The base of this model is the rectangle we obtained from previous sections. The height of the parallelepiped is the elevation averaged from the DEM on the initial blob. Buildings are rendered by putting on the top the texture of the ortho-image and a gray color on the four vertical sides of the buildings. The ground surface is obtained using classical reconstruction with regularisation [1]. The data is the elevation obtained

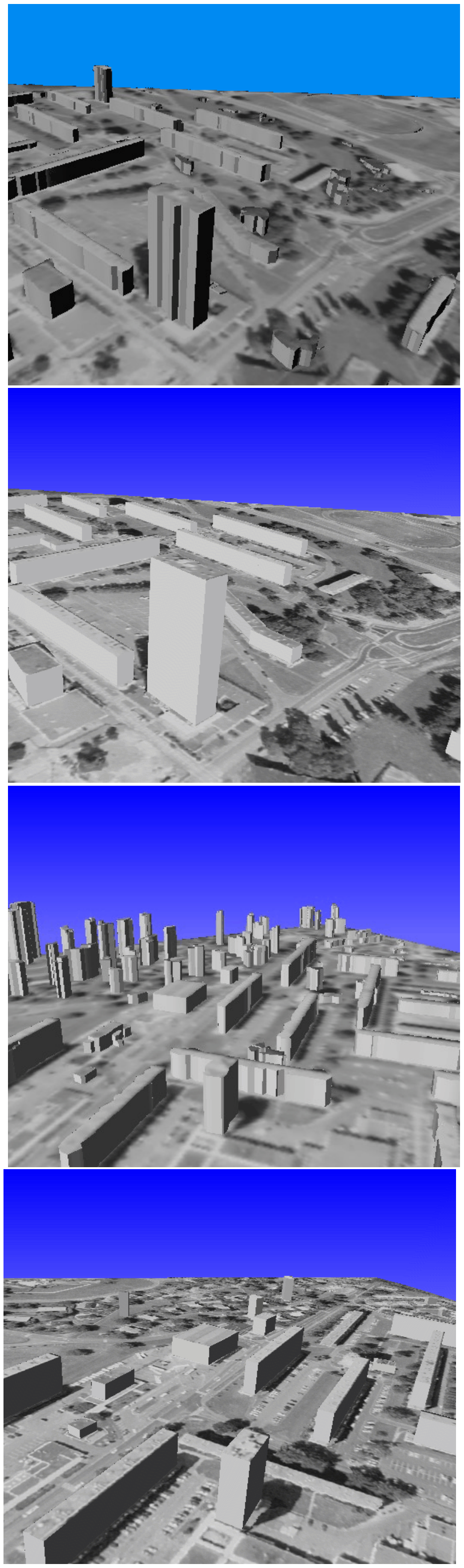

Figure 6. Two different views of 3D reconstruction of the scene without and including rectangle and complex models.

in the DEM, except for location of above-ground blobs where no data is considered. The texture on the surface is the gray level of the ortho-image.

Figure 6 shows two views of 3D reconstruction without and then including our method for rectangle and complex models. The complex building of Fig. 2-left is shown in the middle of Fig. 6-above.

Figure 5 shows how it is important to get a precise location of the rectangle model in order to have a faifthful reconstruction. This figure shows the result of 3D rendering of a rectangular building. We can see that the initial rectangle (as obtained from section 2) gives 3D renderings that are not acceptable views at all both for its shape and location. However, we see in the result image that the initial estimate was close enough to ensure a final result that is perfectly located after energy minimization of section 4. The 3D model of the building is very well inserted in the ground surface and we get a realistic view, as shown in Fig. 6.

6 Conclusion

We described in this paper the processing of the DEM and ortho-image, i.e., the scene vertical view, for the extraction of rectangular buildings as well as buildings which can be decomposed in several rectangles. These automatic processes shorten the operator workload, and compute, to a large extent, the 3D reconstruction of buildings areas. This work has been also applied to dense areas for large cities.

References

[1] A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, 1985.

[2] A. Gruen and R. Nevatia. Special issue on automatic building extraction from aerial images. *Computer Vision and Image Understanding*, 72(2), 1998.

[3] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. In *IJCV*, 1988.

[4] Z. W. Kim, A. Huertas, and R. Nevatia. Automatic description of complex buildings with multiple images. In *WACV*, 2000.

[5] S. C. Lee, A. Huertas, and R. Nevatia. Modeling 3D complex buildings with user assistance. In *WACV'00*.

[6] C. Vestri and F. Duvernay. Improving correlation-based DEMs by image warping and facade correlation. In *CVPR'00*, South Carolina, 2000.

[7] S. Vinson. *3D Reconstruction of geographic sites using variational methods*. PhD thesis, University Paris IX Dauphine, to appear in 2002.

[8] C. Xu and J. L. Prince. Gradient vector flow : A new external force for snakes. In *CVPR'97*.

[9] A. Yuille, P. Hallinan, and D. Cohen. Feature extraction from faces using deformable templates. *IJCV*, 8(2):99–111, Aug. 1992.