Segmentation of Complex Buildings from Aerial Images and 3D Surface Reconstruction

Laurent D. Cohen[1] and Samuel Vinson[1,2]
[1]CEREMADE, UMR CNRS 7534, Université Paris 9 Dauphine, 75775 Paris cedex 16, France
[2]EADS Matra Systèmes & Information
Email: *cohen@ceremade.dauphine.fr, samuel.vinson@iie.cnam.fr*

Abstract

This paper presents a new method for extraction of buildings in aerial images. We first present a method based on rectangular buildings, which are the most common constructions. We then extend this method to more complex shapes by decomposition in a set of rectangles. These rectangles are used to enhance a 3D reconstruction of the digital elevation model (DEM).

Based on stereo data, we use the DEM and the orthoimage for a first segmentation of all areas at elevation above ground. We estimate the rectangle parameters over any given blob and define a criterion for checking the similarity between shape and model. We introduce a new approach for automatic reconstruction of buildings of complex shapes using an iterative splitting of the region until it is covered by a set of rectangles. This automatic process is successfully illustrated on synthetic and real examples. In order to refine location and size of the model, we present a deformable rectangle template.

The final rectangle and complex shape models are used together with elevation to obtain 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, infrastructure of telephony, impact studies, video games,...). For example, authors of [1] have presented automatic methods for terrain modeling (ground, vegetation, buildings...) based on aerial or satellite images. From this modeling and aerial images, the orthoimage, i.e, the vertical view of the scene, is computed (Fig. 2-left). Image resolution is 0.4m per pixel, and size is 2727x2333 pixels.

Automatic building modeling has proven to be a difficult task. There has been active research on this subject these last years [4]. Several works intend to improve building rendering in the Digital Elevation Models (DEM). Lin et al. [8] use perceptual grouping to aggregate building edges. Vestri [10] improves accuracy of DEM by modifying the method of generation by correlation especially on the building frontages.

Lee et al. [7] present a semi-automatic system to generate 3D models with rectilinear hypotheses. This system attempts to minimize the time and the number of user interactions by defining rules to substract or add rectangles to models. Kim et al. [6] 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 use only DEM and orthoimage, and to be less dependent on initial segmentation of above-ground structures. Above-ground structure extraction, like vegetation and constructions, (see Fig. 2-middle) is carried out on the DEM by the algorithms presented in [1, 11], and this paper focuses on the processing after this step is completed. This work was included in building tools for geographic site 3D reconstruction for various types of data including maps (Fig. 1) and aerial images [11].

2 Extraction of Rectangular Buildings

Our first goal is to segment automatically all rectangle buildings from the aerial image. For each blob of Fig. 2-middle, we first find the best matching rectangle and then compute a criteria to test whether the blob is indeed a rectangle.

2.1 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. We make an estimate of the size of the rectangle by assuming the blob is a rectangle. The rectangle

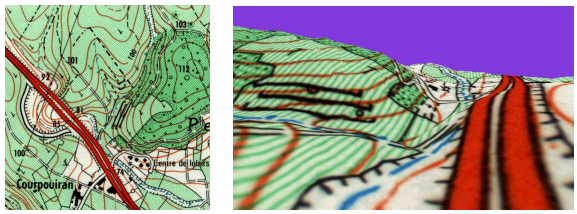

Figure 1: Left: map; right: 3D reconstruction with regularization and constraints based on extracted data from the map (level lines, roads,...).

Figure 2: Segmentation of Rectangular Buildings. Left, the ortho-image, middle, the above-ground areas, and their rectangle approximation on the right, when criteria of similarity by Hausdorff measure is satisfied.

center of mass is the same as the above-ground blob. The orientation θ of the principal axis is computed through the calculation of the second order moments:

$$\tan 2\theta = \frac{2M_{xy}}{M_{xx} - M_{yy}} \text{ and } \mathcal{M} = \begin{pmatrix} M_{xx} & M_{xy} \\ M_{xy} & M_{yy} \end{pmatrix} \quad (1)$$

where the matrix of inertia $\mathcal{M}$ is made of the second order moments and λ_+ and λ_- are its eigenvalues. From the expression of $\mathcal{M}$ as a function of θ, L, ℓ, we can show that

$$\lambda_+ = (L^2 - 1)/12 \quad \text{and} \quad \lambda_- = (\ell^2 - 1)/12. \quad (2)$$

Thus, for a given blob of eigenvalues λ_+, λ_-, the sizes of the best rectangle are obtained by inversion of (2):

$$L = \sqrt{12\lambda_+ + 1} \quad \text{and} \quad l = \sqrt{12\lambda_- + 1} \quad (3)$$

In the case of a square, $M_{xx} = M_{yy}$ and $M_{xy} = 0$, so Eq. 1 gives no value for θ and in fact all directions are principal. However, $L = l$ are known, and the orientation of the square shape is computed by using best correlation between the radial signature of the shape boundary and of a square with same size [11].

2.2 Similarity Criterion

As seen in middle of Fig. 2, some above-ground blobs have not a shape similar to a rectangle. In the following section, we introduce the criterion we selected for the evaluation of our modeling.

The criterion of similarity that we look for, must act as we do when we visually accept or reject an estimate. Consequently, it must be based on a comparison between sets. Based on hand made classification of a large number of regions into three classes (rectangle, non rectangle and questionable), we selected the Hausdorff measure among the criteria we studied [11]. For this criteria, histogram of Hausdorff measure leaded us to a clear threshold value of 0.85 for validating or rejecting our estimate. This means that when our criteria is above that threshold we consider the area as a rectangle and keep its approximation by the rectangle obtained above.

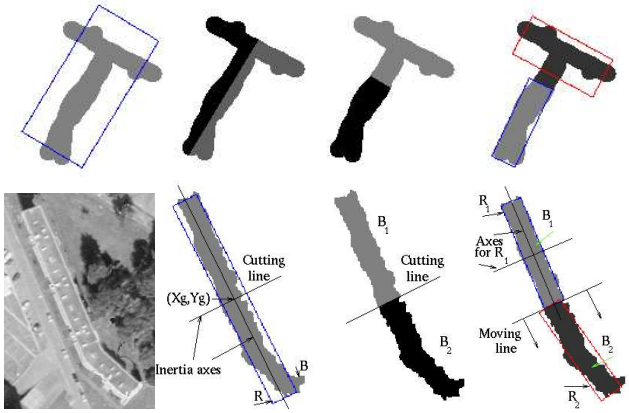

Figure 3: Above from left to right: best rectangle for a synthetic T-shape blob, splitting the blob through the two axes of inertia, best rectangles for the two regions obtained. Below, complex building shape.

The Hausdorff measure $\mathcal{H}$ is a comparison between sets. 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.

Fig. 2-right illustrates 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.

3 Segmentation of Complex Buildings

This rectangular model for buildings is not always sufficient, as shown in the example of Fig. 3. Therefore we have introduced a new method that enables to divide a shape into a set of rectangular shapes. We

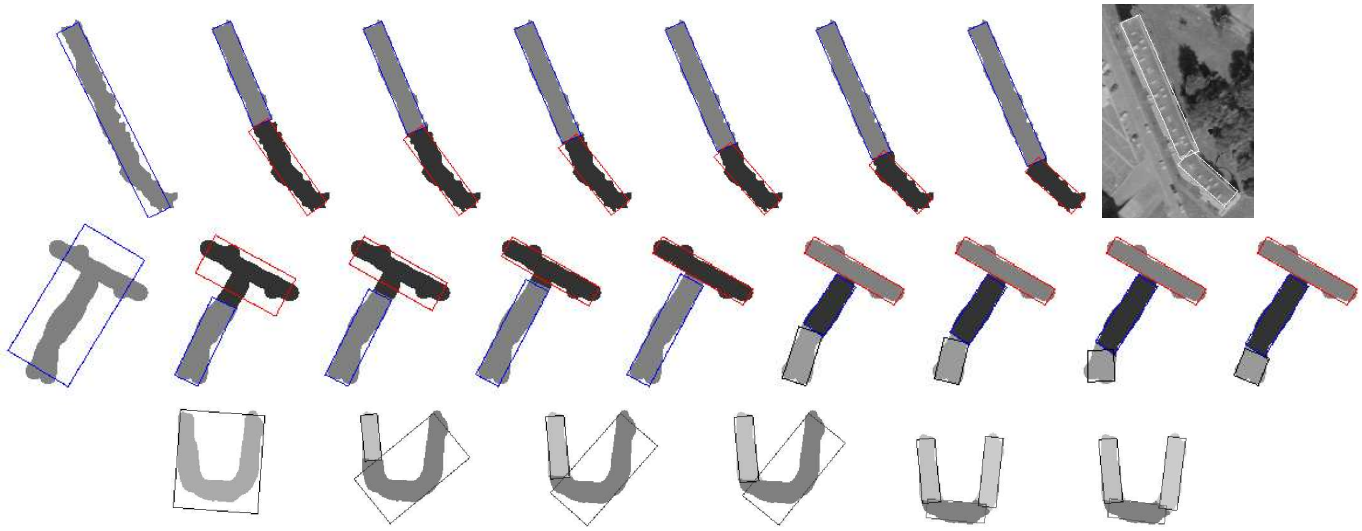

Figure 4: Complex Buildings: for each example, we show from left to right the blob and its best rectangle match, the initial split in two regions, and evolution of this splitting till reaching equilibrium.

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 good rectangle.

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 for the whole blob with cutting line obtained by the smaller eigenvalue axis of inertia. We have some rules in order to have only two connected regions by splitting, as in the example of the U shape of Fig. 4. We denote by $\mathcal{B}$ the complete blob and $\mathcal{B}_1$ and $\mathcal{B}_2$ 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 from the previous section.

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, and we would like to get this matching rectangle as large as possible. A global Hausdorff measure is computed between the union of the two matching rectangles and the complete blob : $\mathcal{H}(R_1 \cup R_2, \mathcal{B})$. 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 estimate. This means the rectangle for which $\mathcal{H}(R_i, \mathcal{B}_i)$ 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 $\mathcal{B}_2$. This means we repeat the two steps starting by splitting $\mathcal{B}_2$ by a line and approach each sub-blob by a rectangle. Further details are given in [11].

Fig. 4 shows the evolution of the process involved for three examples and the final result for decomposition in rectangles. Fig. 5 shows real data results for buildings of complex shapes.

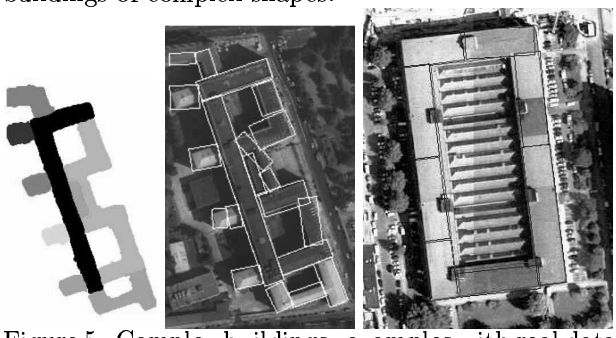

Figure 5: Complex buildings: examples with real data.

4 Refining the rectangular model

Previous computations result in a rectangle or a set of rectangles per above-ground region. When a rectangle is projected onto the orthoimage, its borders may not fully fit the boundary of the building. This inaccuracy is the result of the DEM segmentation process as

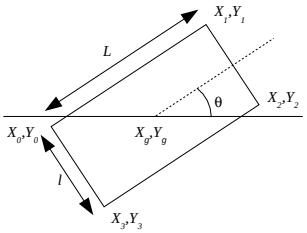

Figure 6: Parametric model of a rectangular building

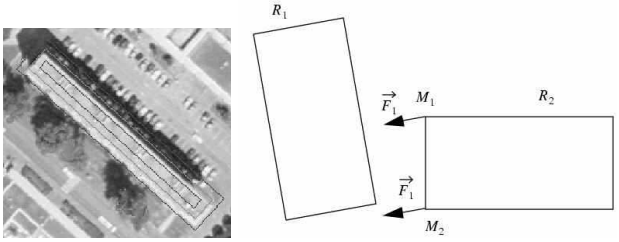

Figure 7: Constraints on the deformable model.

well as the DEM computation itself whose inacurracy is worth a few pixels (up to 10 pixels) [1].

Fua's work concerning netsnakes with hard-constraints [3] results in 3D structure optimization under constraints of horizontality, of verticality and of right angle. However, the method requires the use of several images, whereas we intend to use one image only, i.e. the orthoimage. We use a deformable rectangle template. This is a constrained parametric version of active contours [5], as presented for example by Yuille et al. [13]. 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. Vector of parameters $[a_k] = (X_g, Y_g, \theta, L, \ell)^t$ defines the parametric model (Fig. 6), that evolves by minimizing an energy defined on the four sides of the rectangle:

$$E = E_{01} + E_{12} + E_{23} + E_{30} \quad (4)$$

where the vertices (X_i, Y_i) are indexed from 0 to 3. Each term of energy is derived from a potential P which is computed over the gradient of the orthoimage. This potential must be minimum at the building edges and larger elsewhere.

$$E_{ij} = \int_0^1 P(X_i + \lambda(X_j - X_i), Y_i + \lambda(Y_j - Y_i))\, d\lambda \quad (5)$$

Energy is minimized by means of gradient descent:

$$[a_k]^{n+1} = [a_k]^n - \alpha \nabla E([a_k]^n) \quad (6)$$

where the initial set of parameters is obtained by previous section. As in [2], the descent increment is computed at the beginning of the process such that each parameter move is unitary at the first step, i.e. $\alpha \frac{\partial E}{\partial a_k} = \Delta a_k \approx 1$. For the orientation parameter, the relationship is the following: $\frac{\ell+L}{4}\Delta\theta \approx 1$. Model evolution requires the computation of the energy gradient. The energy partial derivatives computation is the same for each of the four sides of the rectangle.

$$\frac{\partial E_{ij}}{\partial a_k} = \int_0^1 \left(\frac{\partial P(X,Y)}{\partial X}\frac{\partial X}{\partial a_k} + \frac{\partial P(X,Y)}{\partial Y}\frac{\partial Y}{\partial a_k} \right) d\lambda \quad (7)$$

where X and Y are parametric representations of a side of the rectangle as functions of parameters a_k. We refer to [11] for further details.

The potential function P is computed over the orthoimage. It must be minimum at the building borders and high elsewhere. Therefore, as in [5], we compute the opposite of the gradient norm on the orthoimage. Directional gradients are obtained over the image of potential by applying the Gradient Vector Flow [12]. Due to holes in the gradient of some buildings, gradient descent may fail and give a very small rectangle. We thus had to impose some constraints on the model not to go out of the area included between two rectangles inside and outside the initial rectangle (Fig. 7-left). This was done by modifying the potential outside this "permitted area" to attract the rectangle model inside it. Also some extra "internal" forces have to be added in order to deal with simultaneous evolution of a set of rectangles obtained for a complex building. For example, vertices of rectangle R_2 are attracted by a side of R_1 (Fig. 7-right).

Fig. 8 shows results of the deformable model. We can see in this example that the method gives the correct parameters of buildings.

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 and of the above-ground regions. For the buildings, 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 regularization [9]. The data is the elevation obtained in the DEM with holes were left (no data) at the location of above-ground blobs. The

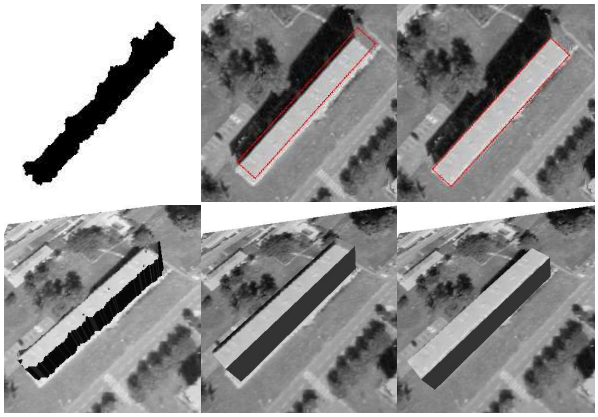

Figure 8: From left to right : initial blob from DEM, initial rectangle, and final result after minimzation. Above the model is superimposed on the ortho-image. Below we show the 3D rendering using the model (see text). Compare the final result with the 3D rendering of the raw blob without rectangle model

texture on the surface is the orthoimage gray level.

We show in Fig. 9 two views obtained from the data of Fig. 2 and from an O-shaped building, including the results of our method for rectangle and complex buildings extraction. The complex building of Fig. 4-above is shown in the middle right of Fig. 9-above.

Fig. 8 shows how it is important to get a precise location of the rectangle model in order to have a faifthful reconstruction. Fig. 8 shows the result of 3D rendering of a rectangular building. We can see that the blob and initial rectangle (as obtained from section 2) give 3D renderings that are not acceptable views at all. 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. 9.

6 Conclusion

We described in this paper the processing of the DEM and orthoimage, i.e., the scene vertical view, for the extraction of rectangular buildings as well as buildings which can be decomposed in several rectangles.

The Hausdorff measure was used to save or reject our estimate. The design of a rectangular parametric model based upon the orthoimage improves the precision of dimensions and localization.

These automatic processes shorten the operator workload, and allow to compute, to a large extent, realistic 3D reconstruction of buildings areas.

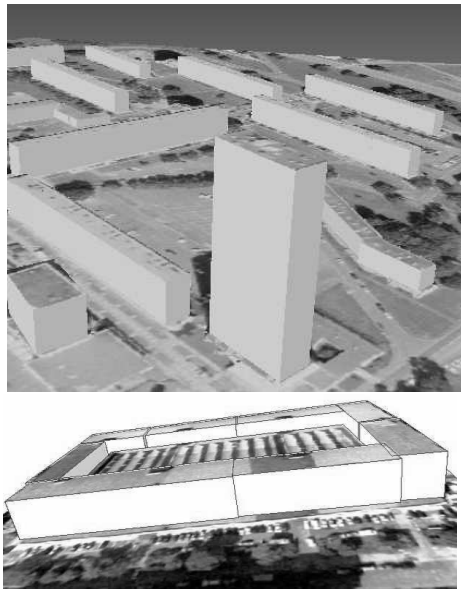

Figure 9: Two different 3D reconstructions of scenes including rectangle and complex models.

References

[1] D. Canu, J.-P. Gambotto, and J. A. Sirat. Reconstruction of building from multiple high resolution images. In *ICIP*, 1996.

[2] Laurent D. Cohen. On active contour models and balloons. *CVGIP:IU*, 53(2):211–218, March 1991.

[3] P. Fua. Fast, accurate and consistent modeling of drainage and surrounding terrain. *International Journal of Computer Vision*, 26(3):1–20, Feb. 1998.

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

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

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

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

[8] C. Lin, A. Huertas, and R. Nevatia. Detection of buildings using perceptual grouping and shadows. In *CVPR'94*.

[9] D. Terzopoulos. The computation of visible-surface representations. *IEEE PAMI*, 10(4):417–438, 1988.

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

[11] Samuel Vinson. *3D Reconstruction of geographic sites using variational methods.* PhD thesis, University Paris 9 Dauphine, April 2002.

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

[13] A. Yuille, D. Cohen, and P. Hallinan. Feature extraction from faces using deformable templates. In *CVPR'89*.