An automatic building reconstruction method: a structural approach using high resolution satellite images

Florent Lafarge, Xavier Descombes, Josiane Zerubia, Marc Pierrot-Deseilligny

HAL Id: hal-00814975

https://inria.hal.science/hal-00814975

Submitted on 18 Apr 2013

AN AUTOMATIC BUILDING RECONSTRUCTION METHOD : A STRUCTURAL APPROACH USING HIGH RESOLUTION SATELLITE IMAGES

Florent Lafarge[1,2], Xavier Descombes[1], Josiane Zerubia[1]   *Marc Pierrot-Deseilligny[2]*

[1]Ariana Reseach Group - INRIA/I3S
2004, routes des Lucioles, BP93
06902 Sophia Antipolis, Cedex France
E-mail=Name.Lastname@inria.fr

[2]Institut Géographique National
2/4 avenue Pasteur
94165 Saint-Mandé, Cedex France
E-mail=Name.Lastname@ign.fr

ABSTRACT

We present an automatic 3D city model of dense urban areas from high resolution satellite data. The proposed method is developed using a structural approach : we construct complex buildings by merging simple parametric models with rectangular ground footprint. To do so, an automatic building extraction method based on marked point processes is used to provide rectangular building footprints. A collection of 3D parametric models is defined in order to be fixed onto these building footprints. A Bayesian framework is then used : we search for the best configuration of models with respect to both a prior knowledge of models and their interactions, and a likelihood which fits the models to the Digital Elavation Model. A simulated annealing scheme allows to find the configuration which maximizes the posterior density of the Bayesian expression.

1. INTRODUCTION

For the last decade, the automatic 3D reconstruction of urban areas has become a topic of interest. Faced with the urbanization development, 3D-models with connected planar facets are used in various applications such as the computing of electromagnetic wave propagation or the creation of virtual realities. Several automatic methods giving satisfactory results, such as perceptual organization [1], parametric models [2] or polyhedral approach [3], have been developed using aerial images.

Nowadays, this problem is tackled by another kind of data : the submetric satellite images. The main advantages of satellite data compared to aerial images are a high swath width and ground coverage. However, such data have a "relatively low" resolution and a "low" signal to noise ratio to deal with 3D reconstruction problems. For example, our satellite data have 4 pixels per square meter contrary to aerial data used in [3] which have about 140 pixels per square meter. Those drawbacks do not allow to robustly use standard methods developed for the aerial image case. The main alternative consists in proposing methods based on an important prior knowledge concerning urban structures.

An automatic building extraction method [4] based on marked point processes is used to provide rectangular building footprints. It consists in extracting the building outlines through a configuration of rectangles from Digital Elevation Models (DEM) which are altimetric descriptions of urban areas. Figure 1 shows the result using a DEM provided from multiple stereo pairs (3-views) of PLEIADES simulations (0.5 meter resolution - B/H=0.2) by the French Geographic Institute (IGN) and computed by an algorithm based on [5]. Our goal is then to construct a 3D city model from the DEM and
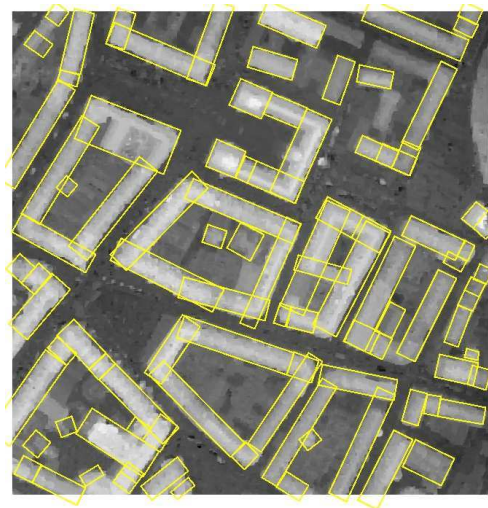

Fig. 1. Building extraction result from a Digital Elevation Model

associated rectangular building footprints. To do so, a structural approach is used : buildings are reconstructed by merging simple urban structures. The last ones are defined through a collection of 3D parametric models with rectangular ground footprint. This approach is preferred since it is less complex, more robust to satellite data and more adapted to rectangular building footprints. Such an approach has been used in [6] based on a restricted grammar of symmetric models in a peri-urban area context from aerial data.

Using a Bayesian framework, the method is based on the definition of a density which contains both a prior knowledge on the buildings, taking into account the interactions existing between neighboring models, and a data term which fits the models to the DEM. A Markov Chain Monte Carlo technique coupled with a simulated annealing is used to find the configuration of models which maximizes this density.

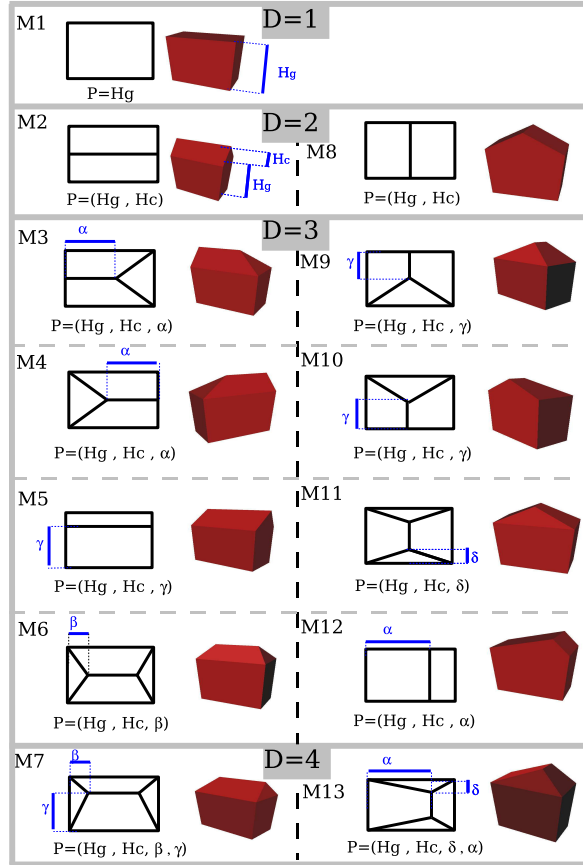## 2. MODEL COLLECTION

2.1. Notations

Let us consider:

- S, a set of sites and $I = \{x(s)/s \in S\}$, a set of intensities defined for a given DEM.

- $\mathcal{R}$, the object space of a rectangle which is defined by five parameters : its center (x_c, y_c) and its length, width and orientation (L, l, ϕ).
- $\mathcal{C} \in \mathcal{R}^N$, the rectangle configuration representing the building footprints associated with I and computed by the method described in [4] (N represents the number of rectangles).
- S_i, the subset of S whose sites are inside the rectangle $i \in \mathcal{C}$.
- $\mathcal{D} = \{x(s) \in I / s \in S_i, i \in \mathcal{C}\}$, the set of data
- $\mathcal{T}$, the state space and $\theta = (\theta_i)_{i \in \mathcal{C}} \in \mathcal{T}$, a configuration of models $\mathcal{M}$.
- f_{θ_i}, the function from S_i to $\mathbb{R}$ which associates the roof altitude of the model defined by θ_i to each site of S_i.

2.2. Choice of the model collection

Figure 2 shows the proposed collection of 3D parametric models with rectangular ground footprint, denoted by $\mathcal{M} = (\mathcal{M}_m)_{m \in [1,13]}$. It is composed of 13 models which are parametrically simple and allow to represent a majority of the common urban structures. For

Fig. 2. Collection of models $\mathcal{M}$ (top views and 3D views) - D, the dimension of the models - P, the parameters of the models

low to represent a majority of the common urban structures. For example, $\mathcal{M}_1$ and $\mathcal{M}_5$ represent respectively flat roofs and dissymmetric two-planes roofs. The 13 proposed models have continuous state spaces with various dimensions (between 1 and 4 without taking into account rectangular base parameters). H_g and H_c are parameters of the roof height which correspond to the getter of roof

height and the roof top height respectively. α, β, γ and δ are parameters of the roof form. They belong to $[0, L]$, $[0, \frac{L}{2}]$, $[0, l]$ and $[0, \frac{l}{2}]$ respectively. More details about these models are available in [7].

3. DENSITY FORMULATION

Let us consider the measurable space $(\mathcal{T}, \mathcal{B}(\mathcal{T}), \mu(.))$ associated to the measure $\mu(.)$ (see [7] for a detailed description of $\mu(.)$). We consider the random variable Θ distributed in $\mathcal{T}$ which follows an unnormalized density $h(.)$ against $\mu(.)$. $h(.)$ is actually the posterior density of a configuration θ of models, given $\mathcal{D}$. In a Bayesian framework, this density can be obtained as follows :

$$h(\theta) = h(\theta/\mathcal{D}) \propto h_p(\theta)\mathcal{L}(\mathcal{D}/\theta) \qquad (1)$$

A requirement is to be able to build both a prior density $h_p(\theta)$ and a likelihood $\mathcal{L}(\mathcal{D}/\theta)$. In the following, these two terms are detailed.

3.1. Likelihood

Let us consider $\mathcal{D}_i$, the partial data of rectangle i defined as $\mathcal{D} = \bigcup_{i \in \mathcal{C}} \mathcal{D}_i$. $\mathcal{L}(\mathcal{D}_i/\theta_i)$ represents the probability of observing $\mathcal{D}_i$ knowing the object θ_i. By considering the hypothesis of conditional independence (it means we disregard the overlapping of rectangles), the likelihood can be expressed as:

$$\mathcal{L}(\mathcal{D}/\theta) = \prod_{i \in \mathcal{C}} \mathcal{L}(\mathcal{D}_i/\theta_i) = \prod_{i \in \mathcal{C}} \exp(-\|f_{\theta_i} - x\|_i) \qquad (2)$$

$\|.\|_i$ is the norm defined from the function space of S_i to $\mathbb{R}$ by :

$$\|f\|_i = \frac{1}{card(S_i)} \sum_{s \in S_i} |f(s)| \qquad (3)$$

So, the likelihood is linked to the Z-error of the L_1 norm between the DEM and the parametric modeling defined by the configuration θ. The L_1 norm is preferred to the L_2 norm since the DEM is neither exact nor accurate. The L_2 norm is too sensitive to the DEM errors.

3.2. Prior density

The prior term allows to favor some configurations and penalize other ones. Some interactions between objects are defined thanks to a neighborhood relationship ν (see Figure 3). The existence of a
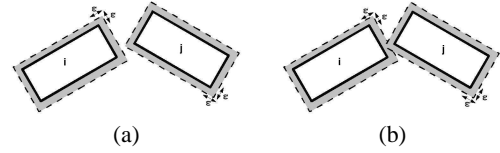
Fig. 3. Neighborhood relationship ν - **(a)** : non neighboring rectangles **(b)** : neighboring rectangles

neighborhood is very important. It allows to consider the problem from a structural point of view by merging models (i.e. by considering a building as an association of rectangles instead of seeing it as a unique rectangle). ϵ defines the neighborhood width. It has been set up to one meter, that is a distance which tolerates small errors concerning the rectangle linking up and is smaller than the average width of a street. The prior density derives from different Gibbs energies developed in the following. It is given by:

$$h_p(\theta) = \exp - [U_h(\theta) + U_r(\theta)] \qquad (4)$$

3.2.1. Getter of roof height adjustment

The getter of roof heights of buildings are dependent on neighboring urban structures. It is important to define an interaction term which favors the getter of roof height alignment between neighboring rectangles. This term has to be:

- attractive for similar getter of roof heights (i.e. with a difference lower than half a floor)
- repulsive for different getter of roof heights (i.e. with a difference between half a floor and one floor)
- neutral for distant getter of roof heights (i.e. with a difference higher than one floor)

To do so, we define the energy term U_h as follows:

$$U_h(\theta) = \sum_{i\nu j} f_h(|H_{g_i} - H_{g_j}|) \tag{5}$$

where f_h is a real value function specified in [7] (see Figure 4), which depends on ω_h, a positive constant potential and H_f, a constant which represents a floor height (in practice, we take $H_f = 3$ meters).
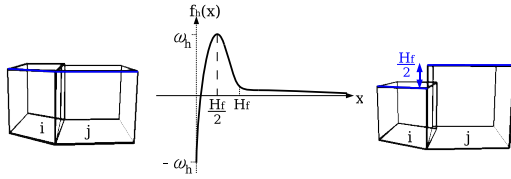

Fig. 4. **left** : favored configuration **center** :function f_h **right** : penalized configuration

3.2.2. Roof top linking up

It is important to develop an interaction which favors roof top linking up between neighboring buildings while respecting the possibility of roof top discontinuity. We propose an interaction which attracts the roof top extremities of neighboring buildings when the roof top height differs from less than a floor height (see Figure 5). The associated energy is modeled as follows:

$$U_r(\theta) = \sum_{i\nu j} \omega_r \, d_c(e_i, e_j) \tag{6}$$

where e_i is the point (in $\mathbb{R}^3$) corresponding to the roof top extremity of the model θ_i and ω_r, a positive constant potential. $d_c(.,.)$ is the function corresponding to the distance related to the $L2$-norm in $\mathbb{R}^3$ if $|(H_{c_i} + H_{c_i}) - (H_{c_j} + H_{c_j})| < H_f$ and 0 if not.
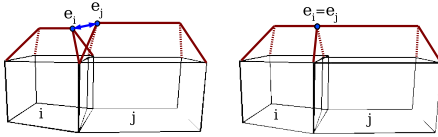
Fig. 5. **left** : penalized configuration **right** : favored configuration

4. OPTIMIZATION

We want to find the configuration which maximizes the density $h(.)$. We search for the Maximum A Posteriori estimator θ_{MAP}:

$$\theta_{MAP} = \arg\max_{\theta} h(\theta) \tag{7}$$

This is a non convex optimization problem in a high and variable dimension space $\mathcal{T}$ since the models of collection $\mathcal{M}$ are defined by a different number of parameters.

A stochastic optimization based on a Reversible Jump Markov Chain Monte Carlo method [8] [9] is well adapted to this problem. It is an optimization method derived from Markov Chain Monte Carlo techniques which allows to deal with variable dimension state spaces. It consists in simulating a discrete Markov Chain $(\Theta_t)_{t\in\mathbb{N}}$ on $\mathcal{T}$ of invariant measure π (specified by the density $h(.)$) which performs small jumps between spaces of variable dimensions. One of the main advantages of such a sampler is that the chain asymptotically converges towards π for all initial configuration Θ_0.

In practice, a simulated annealing is used : the density $h(.)$ is substituted by $h(.)^{\frac{1}{D_t}}$ where D_t is a sequence of temperatures which tends to zero as t tends to infinity. At the beginning of the algorithm (i.e. when the temperature is high), the process is not really selective : it allows to explore the density modes. When the temperature decreases, configurations which have a high density will be favored. Although a logarithmic decrease is necessary to ensure the algorithm convergence, temperature decreases geometrically in practice, in order to reduce the computing time. More details about the optimization process are available in [7].
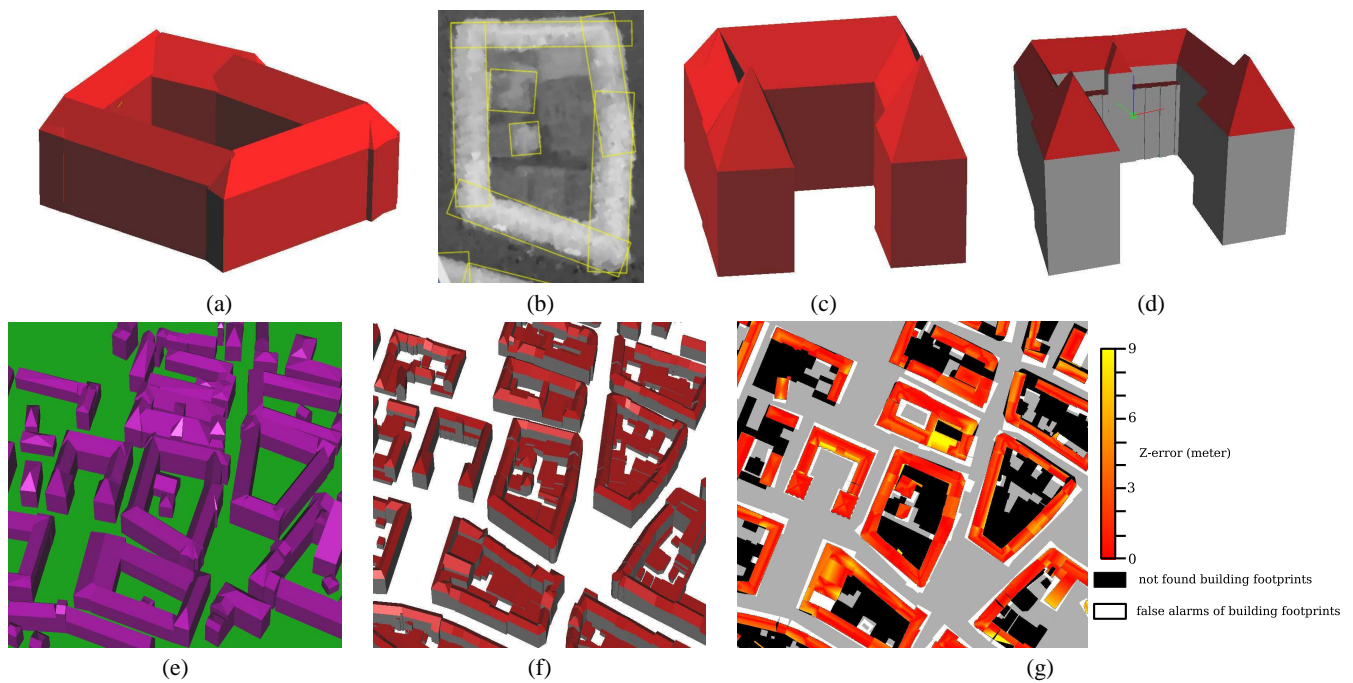
5. RESULTS

In most cases, using energy models implies parameter tuning. Those parameters correspond to weights of the various energy terms ω_h and ω_r, which are chosen by trial and error.

Figure 6-a/c presents two examples of building 3D reconstruction associated respectively with the data (Figure 6-b) and the 3D ground truth provided by the French Geographic Institute (Figure 6-d). The result presented in the figure 6-a is satisfactory compared to the associated data. We can see the importance of the prior knowledge through getter of roof height alignments and roof top linkings up. The proposed modeling in the figure 6-c does not allow to represent some details [7]. However, the global shape is respected and the generalization level is acceptable for satellite data.

Figure 6-e shows the result obtained from Amiens downtown DEM (see Figure 1). It is a satisfactory result with respect to the 3D ground truth (Figure 6-f). The main drawback is the presence of artefacts due to a non optimal rectangle overlapping and roof top linking up impossibility in some specific places. Figure 6-g represents the associated error map which provides three pieces of information. First, it provides the not found areas of the building extraction (in black). They correspond to low flat buildings of inner courtyards that the building extraction method [4] cannot detect. Then, we can see the false alarms of building extraction (in white - rate : 12%), mainly located around the reference building footprint (due to a "drooling" DEM on the building contours which generates wider rectangles). Finally, it provides Z-errors between the reconstruction result and the 3D ground truth (red to yellow). The corresponding Root Mean Square Error (RMSE) of common building footprints is 3.2 meters. This value is satisfactory for a fully automatic method using 2.5 meter Z-resolution DEMs on dense urban areas.

Fig. 6. **(a)-(b)** : example of reconstructed building and the associated data **(c)-(d)** : example of reconstructed building and the associated 3D ground truth ©IGN **(e)-(f)-(g)** : Result on a downtown corresponding to figure 1 - the associated 3D ground truth ©IGN - the associated evaluation map

6. CONCLUSION

Results obtained by the new method proposed in this paper show that the use of a structural approach based on a parametric model collection is well adapted to deal with satellite data in an automatic context. The obtained 3D reconstructions, and especially the roof reconstructions, are satisfactory : a majority of urban structures is close to reality. The proposed prior knowledge allows to compensate for the low quality of data.

However, the main drawback is the presence of artefacts due to a non optimal rectangle overlapping and roof top linking up impossibility at some locations. A solution would consist in using polygonal building footprints by morphing the rectangle configuration into polygon configuration. In the future, we should work on this problem.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] R. Nevatia and K. Price, "Automatic and interactive modeling of buildings in urban environments from aerial images," in *ICIP*, New York, 2002.

[2] U. Weidner and W. Forstner, "Fowards Automatic Building Reconstruction from High Resolution Digital Elevation Models," *Journal of Photogrammetry and Remote Sensing*, vol. 50(4), pp. 38–49, 1995.

[3] C. Baillard and A. Zisserman, "A plane-sweep strategy for the 3D reconstruction of buildings from multiple images," in *19th ISPRS Congress and Exhibition*, Amsterdam, 2000.

[4] M. Ortner, X. Descombes, and J. Zerubia, "Building extraction from Digital Elevation Model.," in *ICASSP*, Honk Kong, Apr 2003.

[5] S. Roy and I.J. Cox, "A maximum-flow formulation of the n-camera stereo correspondence problem," in *ICCV*, Bombay, Jan 1998.

[6] F. Fuchs and H. Le Men, "Efficient subgraph isomorphism with a-priori knowledge. application to building reconstruction for cartography," *Lecture Notes in Computer Science*, , no. 1876, Springer, 2000.

[7] F. Lafarge, X. Descombes, J. Zerubia, and M.P. Deseilligny, "A parametric model for automatic 3D building reconstruction from high resolution satellite images," Research Report 5687, INRIA, september 2005.

[8] P.J. Green, "Reversible Jump Markov Chains Monte Carlo computation and Bayesian model determination," *Biometrika*, vol. 57, pp. 97–109, 1995.

[9] C.P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer-Verlag, New York, 1999.