INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A Parametric Model for Automatic 3D Building Reconstruction from high resolution Satellite Images

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

N° 5687

Septembre 2005

Thème COG

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--5687--FR+ENG

A Parametric Model for Automatic 3D Building Reconstruction from high resolution Satellite Images

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

Thème COG — Systèmes cognitifs
Projets Ariana

Rapport de recherche n° 5687 — Septembre 2005 — 29 pages

Abstract: This report develops a parametric model for automatic 3D building reconstruction based on a Bayesian approach from PLEIADES simulations. High resolution satellite images are a new kind of data to deal with 3D building reconstruction problems. Their "relatively low" resolution and low signal noise ration do not allow to use standard methods developed for the aerial image case. We propose a parametric approach using Digital Elevation Models (DEM) and associated rectangular building footprints. The proposed method is based on a Bayesian approach. A Markov Chain Monte Carlo technique is used to optimize the energy model.

Key-words: 3D building reconstruction, Bayesian approach, Markov Chain Monte Carlo, parametric model.

* Projet Ariana - INRIA/I3S - email=Name.Lastname@inria.fr
† Institut Géographique National - email=Name.Lastname@ign.fr

Un Modèle Paramétrique pour la Reconstruction 3D Automatique de Bâtiments à partir d'Images Satellitaires Haute Résolution

Résumé : Dans ce rapport, nous développons un modèle paramétrique pour la reconstruction automatique de bâtiments en 3D fondé sur une approche bayésienne à partir de simulations PLEIADES. Les images satellitaires haute résolution représentent un nouveau type de données permettant de traiter les problèmes de reconstruction 3D de bâtiments. Leur résolution “relativement basse” et leur faible rapport signal sur bruit pour ce type de problèmes ne permet pas l'utilisation des méthodes standard développées dans le cas des images aériennes. Nous proposons une approche paramétrique utilisant des Modèles Numériques d'Elévation (MNE) et les empreintes de bâtiments associées modélisées par rectangles. La méthode proposée est fondée sur une approche bayésienne. Une technique de type de Monte Carlo par Chaînes de Markov est utilisée afin d'optimiser le modèle énergétique.

Mots-clés : Reconstruction 3D de bâtiments, approche bayésienne, technique de Monte Carlo par Chaînes de Markov, modèle paramétrique.

Contents

Notations

S	a set of sites s
I	a set of intensities x_s associated with S : $I = \{x_s / s \in S\}$
$\mathcal{R}$	object space of a rectangle ($\subset \mathbb{R}^5$)
$\mathcal{C}$	rectangle configuration associated with I ($\in \mathcal{R}^N$)
N	number of rectangles in $\mathcal{C}$
S_i	subset of S whose sites are inside the rectangle $i \in \mathcal{C}$
A_i	roof height space of element $i \in \mathcal{C}$ ($\subset \mathbb{R}^2$)
F_i	roof form space of i ($\subset \mathbb{R}^4$)
T_i	mark space of element i : $T_i = A_i \times F_i$
$\mathcal{T}$	state space : $\mathcal{T} = T_1 \times ... \times T_N$
θ	a element of $\mathcal{T}$: $\theta = (\theta_i)_{i \in \mathcal{C}}$
f_{θ_i}	function from S_i to $\mathbb{R}$ which associates the roof height at each site of S_i
ν	neighborhood relationship defined on $\mathcal{R}$
$h(.)$	a posteriori density on $\mathcal{T}$
$(\Theta_t)_{t \in \mathbb{N}}$	Markov chain on $\mathcal{T}$
$\pi(.)$	target distribution defined on $\mathcal{T}$
D_t	sequence of temperatures
$\mathcal{M}$	collection of simple parametric models
$N_{\mathcal{M}}$	number of models : $N_{\mathcal{M}} = card(\mathcal{M})$
$T_i^{(m)}$	mark space of element i and model $\mathcal{M}_m$
Ψ_{mn}	bijection mapping from model $\mathcal{M}_m$ to $\mathcal{M}_n$
J	jump matrix

1 Introduction

During the last decade, automatic 3D building reconstruction has been a topic of interest. Several automatic methods giving satisfactory results have been developed using aerial images or laser scanning. Aerial images are the most popular data to deal with this problem. Different techniques have been proposed such as perceptual organization [8], parametric models [15] or structural approach [2]. Laser scanning is also very popular since the decrease of acquisition cost. Although laser measurements are very accurate, laser scanning is known to have a low density of points. For instance, we can quote the parametric approach of [6].

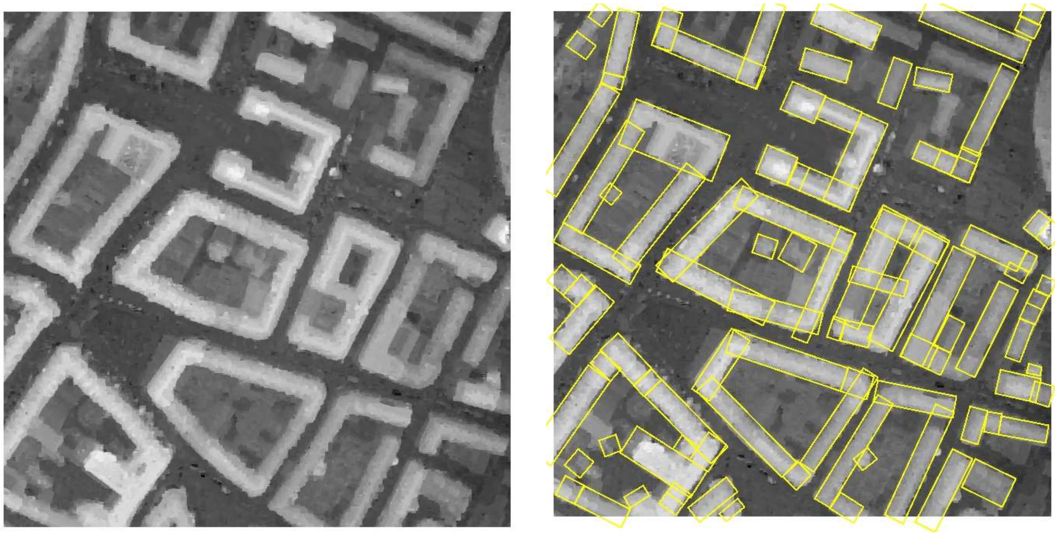
Nowadays, this problem can be tackled by another kind of data : the satellite images. The new satellites own sensors which have sub-metric resolutions. The main advantages of satellite data compared to aerial images are a high swath width and ground coverage. However, such data have a "relatively low" resolution and a low signal noise ratio to deal with 3D building reconstruction. Those drawbacks do not allow to use standard methods developed for the aerial image case and lead us to propose a new method based on important a priori knowledges concerning urban structures.

To do so, a new parametric method for automatic 3D building reconstruction based on a Bayesian approach is developed. This study uses results obtained in the Ariana research group [9] which provide rectangular building footprints based on marked point processes from a Digital Elevation Model (DEM). These results are known to automatically provide efficient building footprints (in 2D). A parametric model is preferred since it is less complex and more robust to satellite data. A priori knowledges of models and their interactions, and a likelihood which fits models to the DEM allow to define an energy. Markov Chain Monte Carlo methods are used to optimize this energy.

2 A Bayesian approach in a continuous state space

2.1 Rectangular building footprint by marked point processes

In this report, we use a building extraction algorithm based on previous work [9], [11]. The authors aimed to extract the building outlines from altimetric descriptions of urban areas. The model [10] is based on marked point processes. These mathematical objects are random variables whose realizations are configurations of geometrical shapes. Objects are rectangles, i.e. simple geometrical forms which allow to efficiently describe building footprints. A data term based on a discontinuity detection is introduced. An a priori term incorporating geometrical interactions between rectangles is also developed in order to regularize the rectangle configuration. An energy is associated to each rectangle configuration, and the global minimum of this energy is then found by applying simulated annealing to a Reversible Jump Monte Carlo Markov Chain sampler. Figure 1 shows a result provided by this method from a DEM based on PLEIADES simulations.



(a) (b)

Figure 1: **(a)** : a Digital Elevation Model (©IGN) **(b)** : extraction result

The results of this method will be used as input data to our problem. Let us consider S, a set of sites and $I = \{x_s/s \in S\}$, a set of associated intensities defined for a given DEM. $\mathcal{R}$ is the object space of a rectangle which is a subset of $\mathbb{R}^5$. So, a rectangle is defined by five parameters : its center (x_c, y_c) and its length, width and orientation (L, l, θ) as shown in the figure 2.

The results of Ortner and al.[10] permit from a DEM defined by $I = \{x_s/s \in S\}$ to know the associated rectangle configuration $\mathcal{C} \in \mathcal{R}^N$ (where N represents the number of rectangles).

Definition 1 *Let us consider $i \in \mathcal{R}$, $\epsilon > 0$. We call :*

- *S_i, the subset of S which sites are inside the rectangle i.*
- *S_i^ϵ, the subset of S which sites are inside the rectangle $i_\epsilon = (x_c, y_c, L + \epsilon, l + \epsilon, \theta)$.*

2.2 The proposed parametric model

The DEM defined by I and the associated rectangle configuration $\mathcal{C}$ are considered as known. The set of data is denoted by $\mathcal{D}$. It is given by :

$$\mathcal{D} = \{x_s \in I/s \in S_i \, , i \in \mathcal{C}\}$$

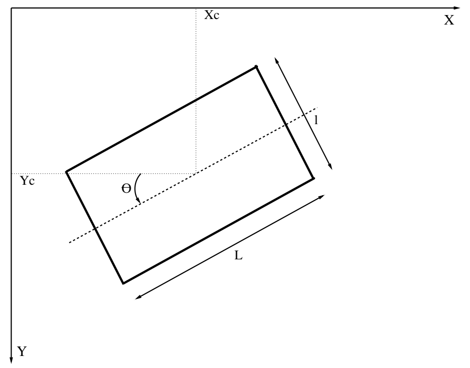


Figure 2: Elements of the object space $\mathcal{R}$

A parametric approach is used. We call T_i, the mark space of element $i \in \mathcal{C}$. $T_i = A_i \times F_i$ is composed of the roof height space A_i and the roof form space F_i.
The roof height space A_i is a compact of $\mathbb{R}^2$ defined by :

$$A_i = \{(H_g, H_c)/H_g \in [H_{gmin}, H_{gmax}], H_c \in [H_{cmin}, H_{cmax}]\}$$

where H_g and H_c represent respectively the getter of roof height and the roof top height of the parametric model (see figure 3).
Roof heights are quite different in urban areas. Therefore, it is important to have well defined roof height spaces A_i. $[H_{gmin}, H_{gmax}]$ and $[H_{cmin}, H_{cmax}]$ must be relevantly fixed for each rectangle. So, a getter of roof height and a roof top height are estimated by using a method we have proposed and which is developed in Annex A. Those estimations allow to center both intervals $[H_{gmin}, H_{gmax}]$ and $[H_{cmin}, H_{cmax}]$ for each rectangle.
The roof form space F_i is a subset of $\mathbb{R}^4$ such that :

$$F_i = \{(a,b,c,d)/(a,b) \in [0,l]^2, (c,d) \in [0,L]^2, a+b \leq l, c+d \leq L and (l-a-b)(L-c-d) = 0\}$$

The term $(l-a-b)(L-c-d) = 0$ constrains the roof top surface to be null (see figure 3). Roof tops with a non-null surface are improbable. Therefore, the roof top of the proposed parametric model is a point or a segment.

Definition 2 *Let us consider $\theta_i \in T_i$, $i \in \mathcal{C}$ and $\eta > 0$. θ_i will be said η-weakly symmetric if $|a-b| < \eta$ **or** $|c-d| < \eta$. θ_i will be said η-strongly symmetric if $|a-b| < \eta$ **and** $|c-d| < \eta$.*

Definition 3 *Let us consider $\theta_i \in T_i$, $i \in \mathcal{C}$. We call f_{θ_i}, the function from S_i to $\mathbb{R}$ which associates the roof height at each site of S_i.*

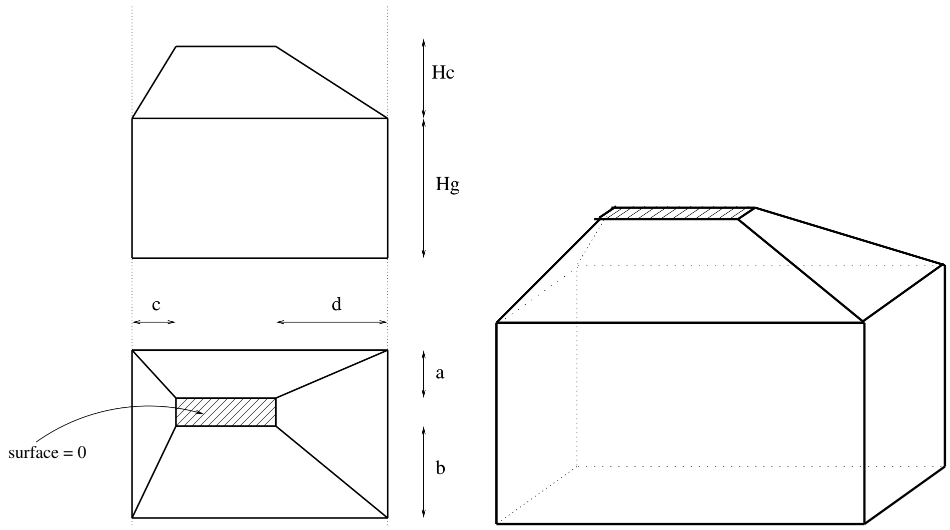


Figure 3: Parametric model of buildings

2.3 Bayesian approach

Here the aim is to find the best configuration $\theta \in \mathcal{T} = T_1 \times ... \times T_N$, where $(1,..,N) = \mathcal{C}$ that follows prefixed criteria. We use a Bayesian approach known to be robust and useful for parameter estimation problems.
It consists in defining the a posteriori density h using Bayes rule. We define the measurable space $(\mathcal{T}, \mathcal{B}(\mathcal{T}), \mu(.))$ associated to the Lebesgue measure $\mu(.)$. We consider the random variable Θ distributed in $\mathcal{T}$ which follows the density h defined as :

$$h(\theta) = \frac{1}{Z} \exp(-U(\theta)) \tag{1}$$

where U is a Gibbs energy and Z a normalizing constant of the density. In most cases, the normalizing constant cannot be computed or estimated since the state space dimension is very large. Then, Z is not evaluated : the density h is not normalized.
The a posteriori density represents the probability of the configuration $\theta \in \mathcal{T}$ knowing the data $\mathcal{D}$. Although it should rigorously be written $h(\theta/\mathcal{D})$, it will be called $h(\theta)$ in this report. By using Bayes rule, we obtain :

$$h(\theta) = \frac{h_p(\theta)\mathcal{L}(\mathcal{D}/\theta)}{g(\mathcal{D})} \propto h_p(\theta)\mathcal{L}(\mathcal{D}/\theta) \tag{2}$$

where $\mathcal{L}(\mathcal{D}/\theta)$ is the likelihood. It represents the probability of observing the data $\mathcal{D}$ knowing the configuration θ. $h_p(\theta)$ is the prior density which allows to regularize the configurations.

The Maximum A Posteriori (MAP) estimator is then used in order to obtain the configuration which maximizes the a posteriori density h :

$$\theta_{MAP} = \arg\max_{\theta} h(\theta) \tag{3}$$

In the next section, we define the likelihood of our model (also called data term) and the prior term (also called regularizing term).

2.4 Likelihood

Let us consider $\mathcal{D}_i$, the partial data of rectangle i defined as $\mathcal{D} = \bigcup_{i \in \mathcal{C}} \mathcal{D}_i$. By considering the hypothesis of conditional independence (it means we disregard the covering of rectangles), the likelihood can be expressed as :

$$\mathcal{L}(\mathcal{D}/\theta) = \prod_{i \in \mathcal{C}} \mathcal{L}(\mathcal{D}_i/\theta_i) \tag{4}$$

$\mathcal{L}(\mathcal{D}_i/\theta_i)$ represents the probability of observing $\mathcal{D}_i$ knowing the configuration θ_i. It is given by :

$$\mathcal{L}(\mathcal{D}_i/\theta_i) = \exp(-\|f_{\theta_i} - x\|_i) \text{ , } i \in \mathcal{C} \tag{5}$$

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

$$\|f\|_i = \frac{1}{card(S_i)} \sum_{s \in S_i} |f(s)| \tag{6}$$

More explicitly, we obtain :

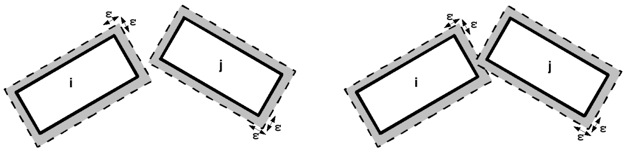
$$\mathcal{L}(\mathcal{D}/\theta) = \exp\left(- \sum_{i \in \mathcal{C}} \frac{1}{card(S_i)} \sum_{s \in S_i} |f_{\theta_i}(s) - x_s| \right) \tag{7}$$

So, the likelihood is linked to the error of the L_1 norm between the DEM and the parametric model of the configuration θ. The L_1 norm is preferred to the quadratic norm since the DEM is neither exact nor accurate. The quadratic norm is too sensitive to the DEM errors.

2.5 Regularizing term

The regularizing term allows to insert constraints about the configuration θ. This term is composed of an a priori related to roof forms and interactions between objects (i.e. energy models of second-order $U(\theta_i, \theta_j)$ where θ_i, θ_j are elements of θ). We want to favor some configurations and penalize other ones.
To do so, a neighborhood between objects of θ must be defined.

Definition 4 *Let us consider θ_i and θ_j, two distinct elements of a configuration θ which are respectively associated with i and $j \in \mathcal{C}$. j is a neighbor of i if and only if $S_i^{\epsilon} \cup S_j^{\epsilon} \neq \varnothing$. We note $i\nu j$, the set of neighbor pairs of the configuration θ.*



(a) (b)

Figure 4: Neighborhood relationship ν - **(a)** : non neighboring rectangles **(b)** : neighboring rectangles

The existence of a neighborhood is very important. It allows to consider the problem at a higher level (i.e. by considering a building or a blockhouse as a collection of rectangles instead of seeing it as a unique rectangle).

The prior density derives from the different energy terms developed in the following. It is given by :

$$h_p(\theta) = \exp - [U_s(\theta) + U_h(\theta) + U_r(\theta)] \tag{8}$$

2.5.1 Roof symmetry

In urban areas, a large majority of buildings have symmetric roofs. Models which are not at least weakly symmetric are improbable models (see figure 5). We aim at favoring the weakly and strongly symmetric models with respect to the other ones.
Let us consider $n_f(\theta)$ and $n_F(\theta)$, the numbers of objects of the configuration θ which are η-weakly and η-strongly symmetric respectively (η is a parameter having a sub-metric value). The constant negative potential ω_f and ω_F are associated with $n_f(\theta)$ and $n_F(\theta)$ respectively. The energy related to the symmetry is then given by :

$$U_s(\theta) = \omega_f n_f(\theta) + \omega_F n_F(\theta) \tag{9}$$

It can be noticed that the strongly symmetric models are more favored than the weakly symmetric ones.

2.5.2 Getter of roof height alignment

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

- attractive for similar getter of roof heights (i.e. with a difference lower than half a floor)

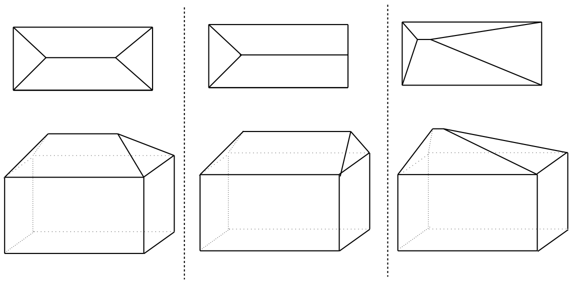


Figure 5: **left** : Roof which is strongly symmetric, **center** : weakly symmetric roof, **right** : improbable model

- 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).

Therefore, the interaction must follow the function f_h (see figure 6), analytically defined as follows :

$$f_h(x) = \begin{cases} \omega_h (1 - \frac{2x}{H_e})^2 & , x \in [0, H_e] \\ \frac{\omega_h}{1+C(x-H_e)^2} & , x \in]H_e, \infty[\end{cases} \tag{10}$$

where H_e is a constant which represents to half a floor height (in practice we take $H_e = 1.5$ meter). ω_h is a constant positive potential. C is a constant $\gg 1$. Then, the associated

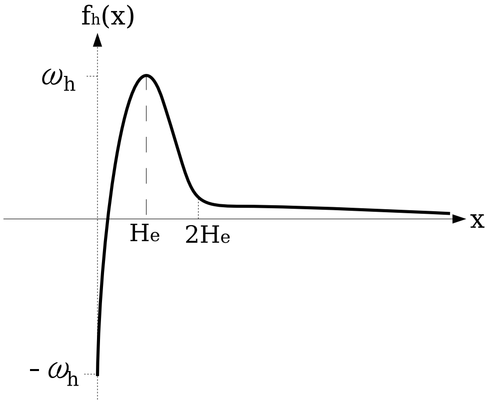


Figure 6: Function f_h

energy is defined as :

$$U_h(\theta) = \sum_{i\nu j} g_h(\theta_i, \theta_j) \tag{11}$$

where g_h is a symmetrical function given by :

$$g_h(\theta_i, \theta_j) = f_h(|H_{g_i} - H_{g_j}|)$$

2.5.3 Roof top linking up

The continuity of roof tops in urban structures is an important point in order to efficiently regularize the model. It is important to develop an interaction which favors the continuity of roof tops between neighboring buildings. More precisely, this term must favor both roof top alignment and roof top linking up as we can see in figures 8 and 7. Figure 8 shows that some configurations are improbable since roof tops exhibit a bad alignment. In the figure 7, we can notice that roof tops must be connected in a single point in order to have a realistic model.

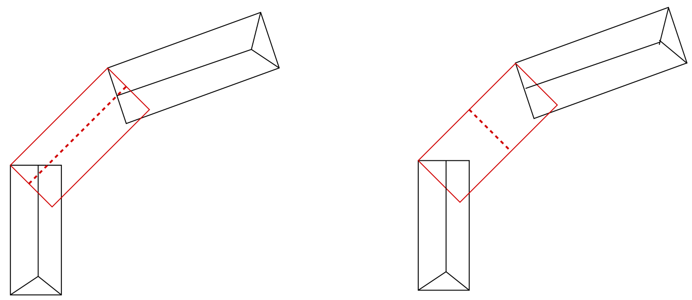


Figure 7: **left** : realistic configuration, **right** : improbable configuration

We propose an interaction which attracts the roof top extremities of neighboring buildings (see figure 9). The associated energy is modeled as follows :

$$U_r(\theta) = \sum_{i\nu j} \omega_r \, d(e_{\theta_i}, e_{\theta_j})^2 \tag{12}$$

where e_{θ_i} is the point (in $\mathbb{R}^3$) corresponding to the roof top extremity of the object θ_i and ω_r a positive constant potential. $d(.,.)$ is the distance related to the $L2$-norm in $\mathbb{R}^3$.

Therefore, this term penalizes configurations for which roof top extremities of neighboring objects are disjoint.

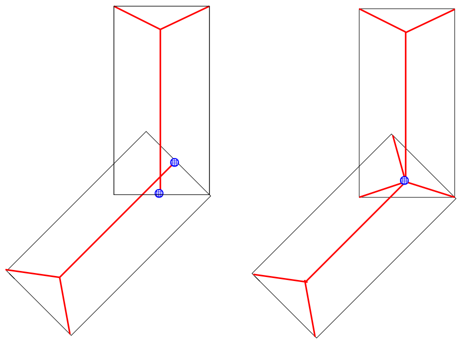


Figure 8: **left** : configuration with roof top extremities which are not linked up, **right** : roof top extremities well connected

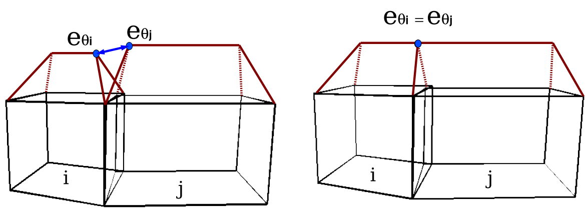


Figure 9: **left** : penalized configuration, **right** : regularized configuration

2.6 Optimization using simulated annealing

We want to find the configuration θ which maximizes the density $h(\theta)$. The Gibbs energy associated with h is not convex. A stochastic optimization based on a Markov Chain Monte Carlo method (MCMC) seems to be well adapted to this problem. It consists in building a discrete Markov chain $(\Theta_t)_{t\in\mathbb{N}}$ from the state space $\mathcal{T}$ which converges to the desired distribution π. This chain must be ergodic so that the distribution asymptotically converges towards π for all Θ_0. The transitions correspond to simple perturbations of configuration θ, they are then simulated in a simple way. The aim is to simulate a chain, from any configuration θ_0, during a number of iterations large enough. The samples will be then distributed by a law close to π.

The Metropolis-Hastings algorithm [5] is used :

- define a matrix of perturbation proposal probability $Q = (q_{ij})_{i,j\in\mathcal{T}}$ where $q_{i.} = q(i \to .)$ is the density of perturbation proposal probability from the state i,
- propose a new state j according to $q(i \to .)$,
- accept the perturbation $i \to j$ with a probability $\alpha_{ij} = \min\{R_{ij}, 1\}$ where R_{ij} is the Metropolis-Hastings acceptance probability defined as follows :

$$R_{ij} = \frac{h(j)q_{ji}}{h(i)q_{ij}}. \tag{13}$$

The Metropolis-Hastings acceptance probability is defined such as h is the stationary measure of the chain $(\Theta_t)_{t\in\mathbb{N}}$.

In practice, a simulated annealing is used : the density $h(.)^{\frac{1}{D_t}}$ is substituted to h 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. However, it is necessary to impose conditions concerning the temperature decay to ensure that the algorithm converges. A logarithmic decrease $D_t = \frac{c}{\log(1+t)}$, where c is a constant high enough to allow the exploration of local minima of energy, permits to be sure the algorithm convergences.
The unnormalized density h has been developed in the previous section. Q needs to be defined. A simple approach consists in taking a symmetric matrix of perturbation proposal probability (i.e. $q_{ij} = q_{ji}$). In that case, none perturbation proposal is favored with respect to other ones : it is the Metropolis algorithm [7]. In this section, we use this algorithm where q_{ij} follows uniform distribution.

2.7 Results

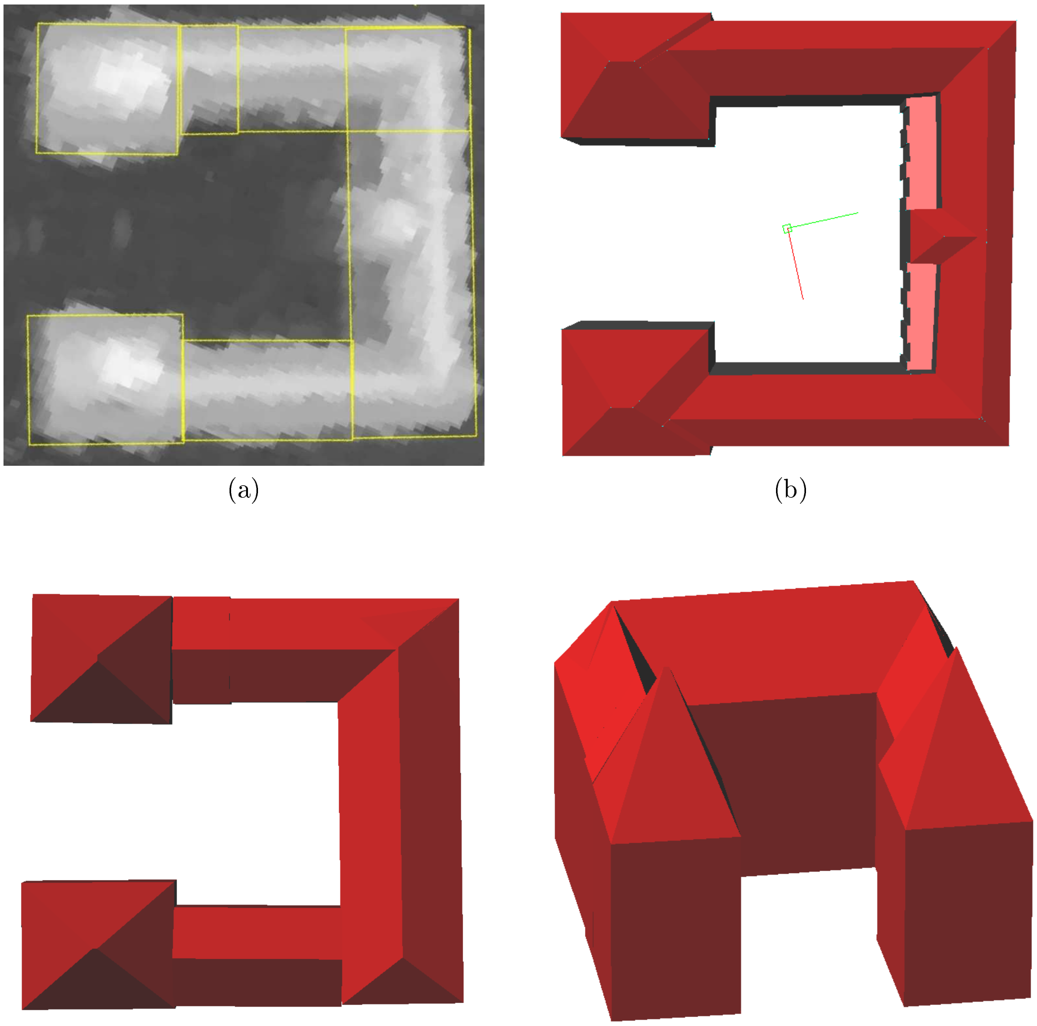
2.7.1 Parameter setting

In most cases, the use of energy models implies parameter tuning. Those parameters mainly correspond to potentials which weight the various energy terms. The likelihood is the reference term. The three other regularizing terms are weighted by one or two values (i.e. ω_f, ω_F, ω_h, ω_a, ω_r). Those four parameters are chosen by trial and error.
It exists other parameters such as ϵ which defines the neighborhood width (see definition 1). The tuning of those parameters is simple because they have a physical sense in our application. For instance, ϵ 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 cooling schedule of the temperature must be relevantly set up. A geometrical decrease has been chosen. It gives satisfactory results and mostly allows to accelerate the optimization process.

2.7.2 Results on PLEIADES simulations

Results have been obtained from DEMs developed at the French Geographic Institute (IGN), using an algorithm based on [13]. We use multiple stereo pairs of PLEIADES simulations (0.5 meter resolution - B/H=0.2) provided by the French Space Agency (CNES). Figure 10-a represents a piece of a DEM showing the City Hall of Amiens (France). The associated rectangle configuration $\mathcal{C}$ appears in yellow. Figure 10-b can be considered as the 3D ground truth of that area. It has been semi-automatically established by the IGN from aerial images (0.2 meter resolution) and ground maps [1].
Figure 10-c represents the top view result. Even if some roof tops are not perfectly linked up (due to the parametric model), the result is good with respect to the ground truth. On the figure 10-d, we can see the importance of a priori knowledges with respect to getter of roof height alignments and roof top linkings up.
We can notice by comparing figures 10-b and 10-c that the rebuilt model seems wider than the real building. It is due to DEMs which have a default, i.e. they are "drooling" on the building contours : the rectangular building footprint is then wider. Figure 11 presents a result on a blockhouse (Amiens - France). Despite a satisfactory reconstruction, this result shows the consequences of rectangle superposition and bad-positioning problems on the 3D modeling. However, a post-processing could allow to compensate this default using a convexity criterion.
Figure 12 shows a result on a larger area which represents Amiens downtown. Figure 12-c 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 [10] cannot detect. Then, we can see the false alarms of building extraction (in white - rate : 12%), mainly located arround the reference building footprints (due to a "drooling" DEM). 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 since we use 2.5 meter Z-resolution DEMs.



(a) (b) (c) (d)

Figure 10: **(a)** :DEM of Amiens City Hall (France) associated with its rectangle configuration **(b)** : 3D ground truth (©IGN - established from aerial images (0.2 meter resolution) and ground maps) **(c)** : result (top view) **(d)** : result (front view)

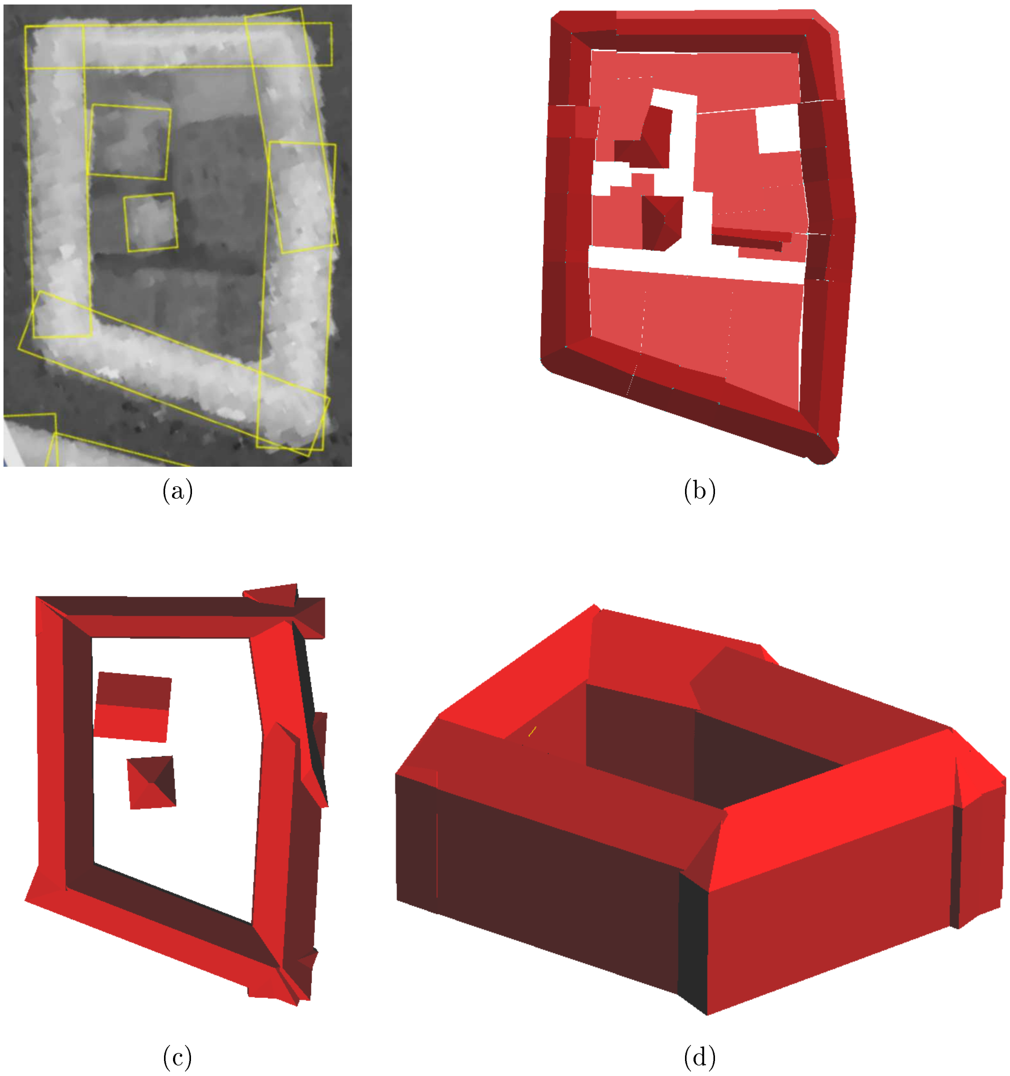


Figure 11: **(a)** : DEM of a Amiens blockhouse (France) associated with its rectangle configuration **(b)** : 3D ground truth (©IGN - established from aerial images (0.2 meter resolution) and ground maps) **(c)** : result (top view) **(d)** : result (front view)

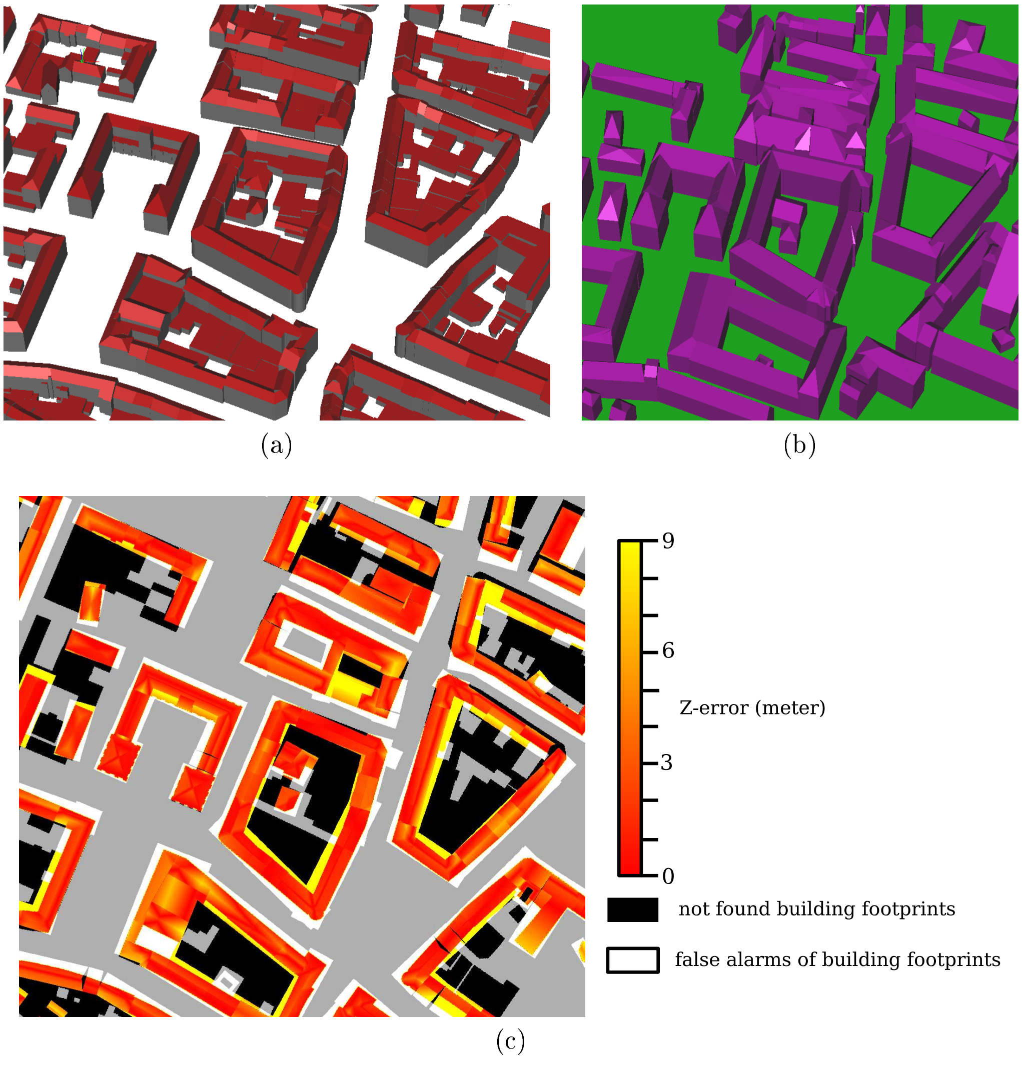


Figure 12: **(a)** : 3D ground truth of Amiens downtown (©IGN - established from aerial images (0.2 meter resolution) and ground maps) **(b)** : result **(c)** : error map

2.7.3 Comments

The results are in general satisfactory. However, the computing time is high. For example, 5 minutes are necessary to obtain the result of Amiens city hall (see figure 10-c / image size $= 158 \times 170$) using a Pentium IV-3Ghz. It is the main problem of the proposed method. However, it is important to notice that a large majority of buildings (close to 90%) have, on the ground truth, a symmetric roof. Our state space $\mathcal{T}$ is a very high dimension space where symmetric models are a particular case. It seems to be relevant to define a new state space which takes into account this symmetry information, to avoid wasting time exploring a space where a large majority of configurations are improbable. This is the topic of the next section.

3 Improvements : towards a variable dimension space

We want to propose a parametric model of buildings which is more relevant than the previous one in order to reduce computing time. It consists in taking into account the high occurrence of symmetric roofs. The parametric model of the previous section is replaced by a collection of simple and relevant models.

3.1 Model collection

Let us consider $\mathcal{M} = (\mathcal{M}_m)_{m \in [1, N_\mathcal{M}]}$, a collection of simple parametric models, where $N_\mathcal{M}$ is the number of models (here, $N_\mathcal{M} = 13$). Figure 13 shows the different models used in this collection (for more details, see annex B). All these models are particular cases of the previous parametric model (see section 2). The state space is then strongly reduced. The new one is composed of relevant models which are parametrically simple and at least weakly-symmetric.

The Bayesian energy model defined in section 2 is used again. Those models have continuous state spaces $T^{(m)}$, $m \in [1, N_\mathcal{M}]$ with various dimensions (between 1 and 4 according to the complexity). The dimension of state space $\mathcal{T}$ is then variable. In this case, the optimization method used in section 2 is not applicable any more. However, it is possible to adapt the MCMC method to the case of a variable dimension state space as shown below.

3.2 RJMCMC

3.2.1 Introduction

The Reversible Jump Markov Chain Monte Carlo method (RJMCMC) has been introduced by Green [4]. He has extended the formalism introduced by Hastings [5] by allowing the Markov chain to realize small jumps between spaces of variable dimensions respecting the reversibility assumption of the chain. Propositions are based on “small jumps” which means only one object of the global configuration will be concerned by a new proposition. This algorithm builds a Markov chain $(\Theta_t)_{t \in \mathbb{N}}$ using :

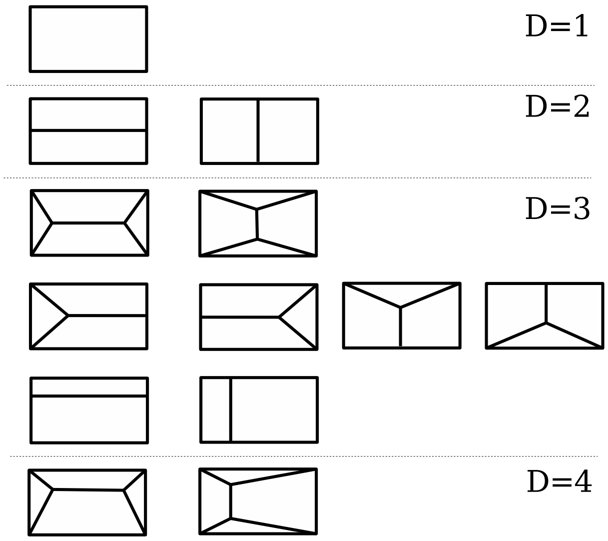


Figure 13: Collection of parametric models (top view) - D is the dimension of the parametric models

- the objective law $\pi(.)$ of the chain defined on $\mathcal{T}$ and known up to a normalizing constant ($h(.)$ is the unnormalized density of $\pi(.)$).
- proposition kernels $Q_k(.,.)$ defined on $\mathcal{T} \times \mathcal{B}(\mathcal{T})$.
- a symmetric measure $\psi_k(.,.)$ associated with $Q_k(.,.)$ and defined from $\mathcal{T} \times \mathcal{T}$ such as the joint measure $\pi(.)Q_k(.,.)$ is absolutely continuous with respect to $\psi_k(.,.)$. This assumption allows to ensure the existence of the Radon-Nikodym derivative $f_k(.,.)$ defined as follows :
$$f_k(x,y) = \frac{\pi(dx)Q_k(x,dy)}{\psi_k(dx,dy)} \tag{14}$$

The probability of accepting the proposition is $\min(1, R_k)$ where R_k is the Green ratio given by :

$$R_k(x,y) = \frac{f_k(y,x)}{f_k(x,y)} = \frac{\pi(dy)Q_k(y,dx)}{\pi(dx)Q_k(x,dy)} \tag{15}$$

The reversibility assumption is ensured due to the symmetry of the measure ψ_k. RJMCMC technique is especially well adapted to point processes as shown in [3].

3.2.2 Jumps between models

Let us consider two models $\mathcal{M}_m$ and $\mathcal{M}_n$ and a move (or jump) from $x = (m, \theta_m)$ to $y = (n, \theta_n)$. The idea of Green is to create a bijection between the spaces $T^{(m)}$ and $T^{(n)}$. θ_m is completed by a simulation $u_{mn} \sim \varphi_{mn}(u_{mn})$ into (θ_m, u_{mn}), and θ_n by $v_{nm} \sim \varphi_{nm}(v_{nm})$ into (θ_n, v_{mn}) such as the mapping Ψ_{mn} between (θ_m, u_{mn}) and (θ_n, v_{mn}) is a bijection :

$$(\theta_n, v_{mn}) = \Psi_{mn}(\theta_m, u_{mn}) \tag{16}$$

The probability of acceptance for the move from $\mathcal{M}_m$ to $\mathcal{M}_n$ is then :

$$\min\left(\frac{\pi(n,\theta_n)}{\pi(m,\theta_m)}\frac{J_{nm}\varphi_{nm}(v_{nm})}{J_{mn}\varphi_{mn}(u_{mn})}\left|\frac{\partial\Psi_{mn}(\theta_m,u_{mn})}{\partial(\theta_m,u_{mn})}\right|, 1\right) \tag{17}$$

where J_{mn} is the probability of choosing a jump to $\mathcal{M}_n$ while in $\mathcal{M}_m$ and φ_{mn}, the density of u_{mn} (here, the u_{mn} follow uniform distributions).
Finding the bijection Ψ_{mn} is a difficult problem when $\mathcal{M}_m$ and $\mathcal{M}_n$ differ from many dimensions [12]. In practice, the jumps are limited to moves from $\mathcal{M}_m$ to models with close dimensions. It is the case in our problem : the $N_{\mathcal{M}}$ prefixed models differ from few dimensions. The jumps from any model to any other one are then allowed.
For instance, we consider the jump from $\mathcal{M}_2$ to $\mathcal{M}_4$ (see figure 14). We move from $\theta_2 = (H_g, H_c) \in T_2$ to $\theta_4 = (H_g, H_c, \alpha) \in T_4$ ($\alpha \in [0, L]$). We then generate $u_{2\,4}$ according to a uniform distribution on $[0, L]$. So, we have $\theta_4 = \Psi_{2\,4}(\theta_2, u_{2\,4}) = Id(\theta_2, u_{2\,4})$ and $\left|\frac{\partial\Psi_{2\,4}(\theta_2,u_{2\,4})}{\partial(\theta_2,u_{2\,4})}\right| = 1$.
Now, let us consider a more complex example : a jump from $\mathcal{M}_6$ to $\mathcal{M}_{13}$. We move from $\theta_6 = (H_g, H_c, \beta) \in T_6$ to $\theta_{13} = (H_g, H_c, \delta, \alpha) \in T_{13}$. It exits a bijection between β and δ : $\delta = \frac{l}{L}\beta$. In fact, β, δ and α are variables belonging to $[0, \frac{L}{2}]$, $[0, \frac{l}{2}]$ and $[0, L]$ respectively. Then, in generating $u_{6\,13}$ according to a uniform distribution on $[0, L]$, $\theta_{13} = \Psi_{6\,13}(\theta_6, u_{6\,13}) = A(\theta_6, u_{6\,13})$ where $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{l}{L} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. We obtain $\left|\frac{\partial\Psi_{6\,13}(\theta_6,u_{6\,13})}{\partial(\theta_6,u_{6\,13})}\right| = \frac{l}{L}$.
Thanks to the reversible assumption, it is important to notice that :

$$R_{nm}(\theta_n, \theta_m) = [R_{mn}(\theta_m, \theta_n)]^{-1} \tag{18}$$

The Green algorithm can be resumed as follows :

Algorithm 1 *At iteration t, if* $x^{(t)} = (m, \theta_m^{(t)})$,

- *select model* $\mathcal{M}_n$ *with probability* s_{mn}
- *generate* $u_{mn} \sim \varphi_{mn}(u)$
- *set* $(\theta_n, v_{mn}) = \Psi_{mn}(\theta_m^{(t)}, u_{mn})$

- *take* $\theta_n^{(t)} = \theta_n$ *with probability*

$$\min\left(\frac{\pi(n,\theta_n)}{\pi(m,\theta_m^{(t)})}\frac{J_{nm}\varphi_{nm}(v_{nm})}{J_{mn}\varphi_{mn}(u_{mn})}\left|\frac{\partial\Psi_{mn}(\theta_m^{(t)},u_{mn})}{\partial(\theta_m^{(t)},u_{mn})}\right|,1\right)$$

and take $\theta_m^{(t+1)} = \theta_m^{(t)}$ *otherwise.*

3.2.3 Adaptive cooling schedule

The choice of the initial temperature and of the decreasing schedule of the temperature is an important parameter in simulated annealing. It exists a temperature, called the critical temperature, for which the process chooses its energy shaft. Time must be spent at this temperature in order to select the correct minima of energy. Several adaptive cooling schedules have been proposed as for instance in [9] or [14].
The idea consists in searching for the critical temperature by detecting high energy variations during some iterations. We then slow down the temperature decrease in order to have a detailed exploration. Once the shaft is selected, we accelerate the temperature decrease. In practice, different geometric decreases are used.

3.2.4 Adaptive jump matrix

The jump matrix $J = (J_{mn})_{m,n\in[1,N_{\mathcal{M}}]}$ denotes the probability that a proposed jump to model $\mathcal{M}_n$ is attempted starting from anywhere in T_m. The convergence speed of the algorithm directly depends on J. Proposing relevant jumps will favor the algorithm convergence. J must be efficiently initialized in order to get a faster convergence. In figure 14, we can see that some models are close in terms of roof structures : we call them neighboring models. Jumps between neighboring models are then favored in the initial matrix J by imposing $J_{mn} = 2J_{ml}$ with $\mathcal{M}_n$ being a neighboring model of $\mathcal{M}_m$ ($\mathcal{M}_l$ being some model).
Moreover, matrix J must be able to evolve during iterations with respect to the previous accepted and rejected propositions. More precisely, all N_a iterations, we update J with respect to rates of accepted propositions : the higher the acceptation rate of a jump, the higher the jump coefficient of J.
Let us consider p_{mn}, the quotient of the accepted proposition number by the refused proposition number for the jumps from $\mathcal{M}_m$ to $\mathcal{M}_n$ established during $[t, t+N_a]$. N_a must be high so that p_{mn} are robust. At iteration $t + N_a$, s_{mn} is updated as follows :

$$J_{mn} \leftarrow \rho J_{mn} + (1-\rho)\frac{p_{mn}}{\sum_n p_{mn}} \tag{19}$$

where ρ is a coefficient belonging to $]0,1]$. The fact that $\rho \neq 0$ ensures the Markov chain irreducibility.

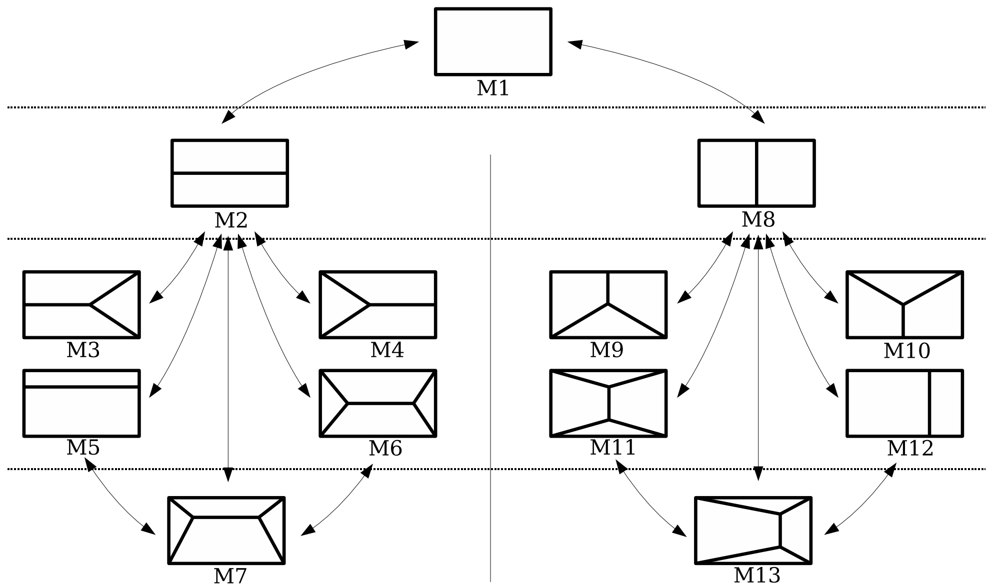


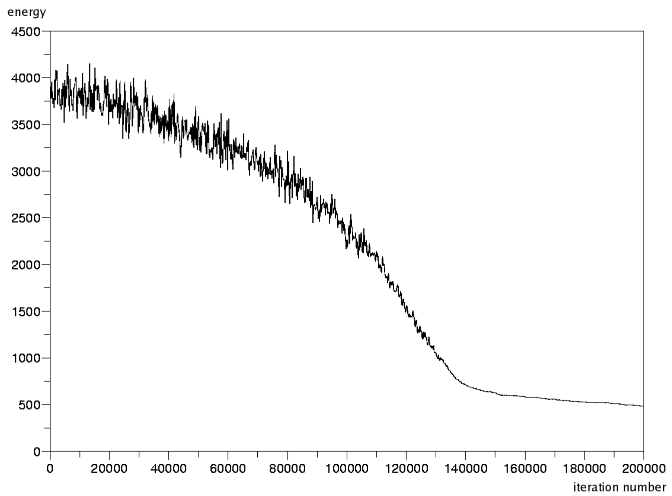
Figure 14: Jumps between neighboring models

3.3 Results

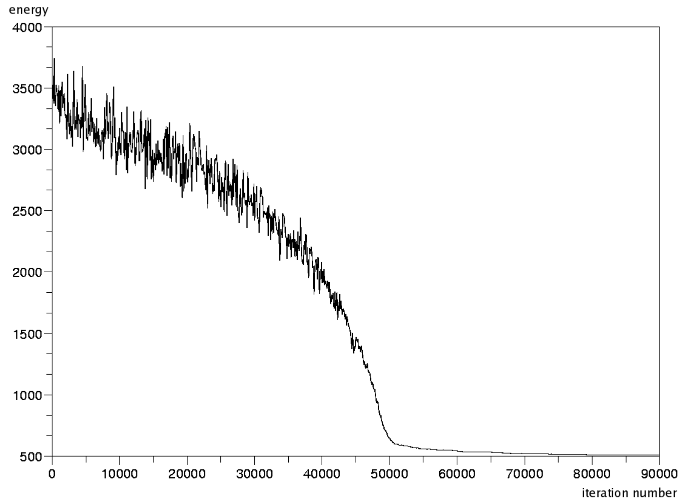
The results are identical to those of section 2. The model collection allows a correct description of urban scenes. The main advantage is the computing time which is shorter. Computing time is improved of a factor 2. Although the adaptative jump matrix introduces new parameters ρ and N_a to be set up, it allows a better convergence. In practice, we take $\rho = 0.8$ and $N_a = 10,000$.
Figure 15 shows the convergence of both algorithms. We can see that using RJMCMC algorithm allows a faster convergence compared to Metropolis algorithm.

4 Conclusion

Results obtained by the proposed approach show that the use of a parametric model is well adapted to deal with high resolution satellite images in a fully automatic context. The 3D building reconstruction, and especially the roof reconstruction, is close to reality : the RMSE between the reconstruction result and the 3D ground truth is 3.2 meters. It is a satisfactory result with respect to a 2.5 meter Z-resolution DEM. Computing times have been improved by using a model collection (that reduces space data) and a RJMCMC optimization. However, it seems to be necessary to correct two drawbacks : artifacts due to a non optimal rectangle overlapping and roof top linking up impossibility at some areas. Both



(a)



(b)

Figure 15: Energy convergence - **(a)** Metropolis (section 2) **(b)** RJMCMC (section 3)

defaults should be eliminated in the future by using post-processing based on improvements of rectangular building footprints.

Acknowledgments

This first author would like to thank the French Geographic Institute (IGN) and the French Space Agency (CNES) for partial financial support during his PhD.

References

[1] D. Flamanc and G. Maillet. Evaluation of 3D city model production from PLEIADES HR satellite images and 2D ground maps. In *URBAN*, Tempe, U.S., 2005.

[2] F. Fuchs and H. Le Men. Efficient subgraph isomorphism with a-priori knowledge. application to building reconstruction for cartography. *Lecture Notes in Computer Science*, (1876), Springer, 2000.

[3] C.J. Geyer and J. Moller. Simulation and likehood inference for spatial point processes. *Scandinavian Journal of Statistics*, Series B(21):359–373, 1994.

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

[5] W.K. Hastings. Monte Carlo sampling using Markov chains and their applications. *Biometrica*, 57(1):97–109, 1970.

[6] H. Maas. Fast determination of parametric house models from dense airborne laser-scanner data. In *ISPRS workshop on Mobile Mapping Technology*, 1999.

[7] M. Metropolis, A.W. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087–1092, 1953.

[8] R. Nevatia and K. Price. Automatic and interactive modeling of buildings in urban environments from aerial images. In *ICIP*, pages 525–528, New York, 2002.

[9] M. Ortner. *Processus ponctuels marqués pour l'extraction automatique de caricatures de bâtiments à partir de modèles numériques d'élévation*. PhD thesis, University of Nice-Sophia Antipolis, 2004.

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

[11] M. Ortner, X. Descombes, and J. Zerubia. Building outline extraction from Digital Elevation Models using marked point processes. *To appear in International Journal of Computer Vision*, 2005.

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

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

[14] R. Stoica, P. Gregori, and J. Mateu. Simulated annealing and object point processes : tools for analysis of spatial patterns. Technical Report 69, University Jaume I, Castellon, Spain, 2004.

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

Annex A : roof height estimations

Here, we describe a method to estimate the roof top height and the getter of roof height from a rectangular building footprint. In fact, we estimate those heights for each rectangle of the footprint by computing means on rectangle masks.

Getter of roof height estimation

The first step consists in selecting an area of the rectangle which is close to the getter of roof by defining a mask (see figure 16). The mask width is proportional to the rectangle width by a factor $\frac{1}{5}$. Then, we select pixels of DEM which belong to the mask and we rank them by increasing order. If we call N_g the number of selected pixels, we compute the mean of pixel values which are ranked between $0.3N_g$ and $0.7N_g$. The result gives us the getter of roof height estimation of the considered rectangle.

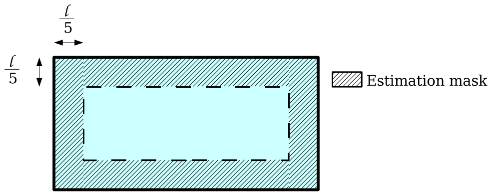


Figure 16: Mask of the getter of roof height estimation

Roof top height estimation

The principle is similar to the Getter of roof height estimation. We define a mask close to the roof top (see figure 17).The mask width is proportional to the rectangle width by a factor $\frac{1}{4}$. We rank as previously the N_c pixels of the DEM which belong to the mask. We compute the mean of pixel values which are ranked between $0.85N_c$ and $0.95N_c$ in order to obtain the roof top height estimation of the rectangle.

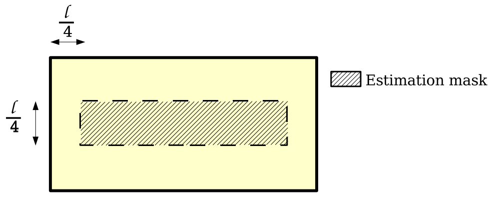


Figure 17: Mask of the top roof height estimation

Annex B : collection of parametric models

We detail a collection of simple parametric models used in section 3. Those models have continuous state spaces with various dimensions (between 1 and 4). Figure 18 presents the 13 models. H_g and H_c are the 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 the parameters of the roof form. They belong to $[0, L]$, $[0, \frac{L}{2}]$, $[0, l]$ and $[0, \frac{l}{2}]$ respectively.

For instance, model $\mathcal{M}_1$ is a flat roof : it is a very simple model which owns an unique parameter (H_g). In the contrary, Model $\mathcal{M}_{13}$ is a more complex roof. It owns 4 parameters $(H_g, H_c, \delta, \alpha)$.
Models $\mathcal{M}_2$ to $\mathcal{M}_7$ have a roof top which is oriented following the rectangle length. In the contrary, the roof top of models $\mathcal{M}_8$ to $\mathcal{M}_{13}$ are oriented following the rectangle width.
Five models are strongly symmetric ($\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_6$, $\mathcal{M}_8$ and $\mathcal{M}_{11}$). The other ones are weakly symmetric.

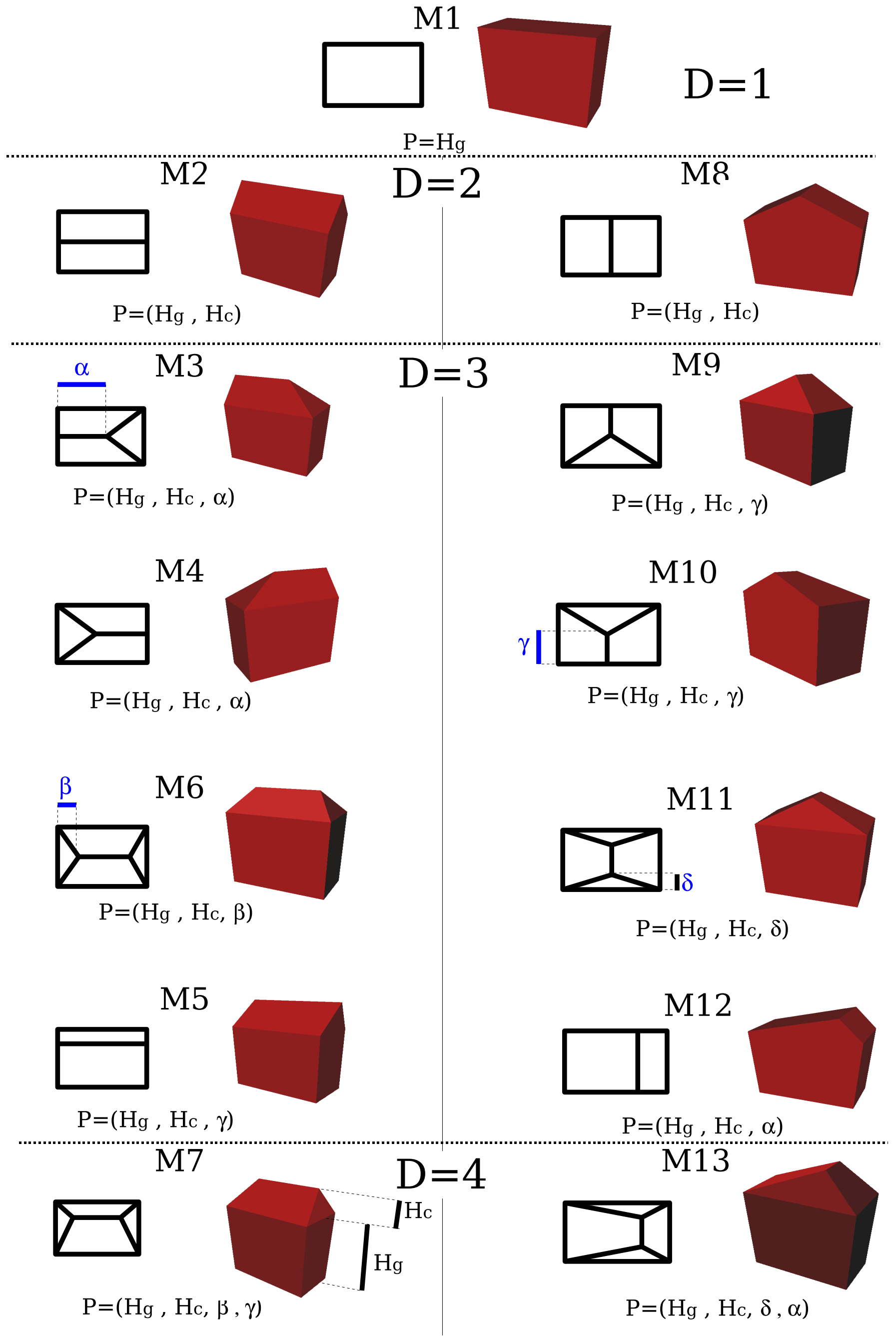


Figure 18: Collection of models $\mathcal{M}$ (top view and 3D view) - D, the dimension of the models - P, the parameters of the models

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399