

Building reconstruction from a single DEM

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

▶ **To cite this version:**

Florent Lafarge, Xavier Descombes, Josiane Zerubia, Marc Pierrot-Deseilligny. Building reconstruction from a single DEM. Computer Vision and Pattern Recognition (CVPR), Jun 2008, Anchorage, United States. hal-00781733

HAL Id: hal-00781733
https://inria.hal.science/hal-00781733v1

Submitted on 28 Jan 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Building reconstruction from a single DEM

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

[1] Ariana Research Group - INRIA
Sophia-Antipolis, France
Firstname.Name@inria.fr

[2] French Mapping Agency (IGN)
Saint-Mandé, France
Firstname.Name@ign.fr

Abstract

We present a new approach for building reconstruction from a single Digital Elevation Model (DEM). It treats buildings as an assemblage of simple urban structures extracted from a library of 3D parametric blocks (like a LEGO®set). This method works on various data resolutions such as 0.7 *m satellite and* 0.1 *m aerial DEMs and allows us to obtain 3D representations with various levels of detail. First, the 2D supports of the urban structures are extracted either interactively or automatically. Then, 3D blocks are placed on the 2D supports using a Gibbs model. A Bayesian decision finds the optimal configuration of 3D blocks using a RJMCMC sampler. Experimental results on complex buildings and dense urban areas are presented using data at various resolutions*[1].

1. Introduction

Three dimensional models of urban areas are very useful for many applications such as urban planning, radiowave reachability tests for wireless communications, disaster recovery or computer games.

Problem statement Many automatic methods have been proposed. It is difficult to compare these methods efficiently since they have been developed in different contexts (kinds of data, types of reconstructed buildings, level of user interactivity, *etc*) and use different evaluation criteria. Multiple view images are the most common inputs. Scholze *et al.* extracted 3D-lines and grouped them into faces which allow the building reconstruction through a semantic interpretation [12]. Rooftop hypotheses were generated from 3D-lines and junction information by Kim *et al.*[5]. Baillard *et al.* present a method based on planar facet hypothesis which can be generated from single 3D-lines [1]. Mueller *et al.* developed a procedural model based on a shape grammar [9]. One of the most efficient methods uses a hierarchical approach which combines three different levels of detail [8]. Laser scans are also popular inputs due to the measurement accuracy. Fruech *et al.*[3] uses laser scans to model buildings with a detailed reconstruction of the facades. These methods provide convincing 3D-models using aerial images, ground views or laser scans. However, most of them have been developed using a specific kind of data and cannot easily be adapted if the image characteristics change (resolution, SNR, *etc*). We propose a new method adapted to varying data resolution.

Global strategy There are two main families of approaches in 3D building reconstruction. Generic representations are theoretically able to reconstruct any shape of building through connected planar facets, but they demand high resolution data. Parametric representations are known to be robust with respect to data quality [14], but these reconstructions are limited - most parametric representations consider a symmetric two-plan roof reconstruction.
In this paper, we use an approach which is halfway between generic and parametric representations : the structural approach. It consists in reconstructing buildings by assembling simple urban structures extracted from a library of 3D parametric blocks, as a LEGO®game (see Figure 1). This approach is particularly interesting since it combines the advantages of generic and parametric representations:
1- the robustness of parametric approaches is preserved since the library objects are defined by parameter sets,

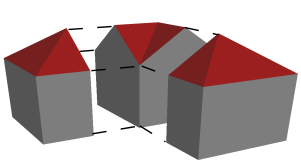


Figure 1. Principle of the structural approach.

[1]The first author thanks the French Mapping Agency (IGN) and the French Space Agency (CNES) for partial financial support during his PhD. The authors thank the CNES for providing satellite images.

2- an efficient library permits modeling a large range of buildings. It is even possible to reconstruct buildings such as curved roof structures that some generic models cannot construct,
3- assembly of urban structures is particularly adapted to multiresolution data and allows us to obtain both basic roof shapes from 70 cm resolution satellite images or details including superstructures (chimneys, dormer-windows,*etc*) from 10 cm resolution aerial images.
This approach is based on important prior knowledge concerning urban structures and their assembly. It is necessary to define the interactions between blocks correctly to have a convincing modeling without artefacts. A stochastic framework is especially well adapted to introduce such knowledge. Lafarge *et al.* proposed a similar approach [7]. However, that work was limited and suffered from many drawbacks: 1-generation of many artefacts, 2-lack of 3D-modeling generality, 3-tuning of many parameters, 4-long computation times and 5-restriction to satellite data.
The input was a single Digital Elevation Model (DEM) which is well adapted to global geometric descriptions. First, the 2D-supports of the urban structures were extracted either interactively or automatically using Ortner *et al.*'s method [10]. Then, 3D blocks are positioned on the 2D supports using a Gibbs model in Section 3. A Bayesian decision finds the optimal configuration of 3D blocks using a RJMCMC sampler. Experimental results on complex buildings and dense urban areas are shown using data at various resolutions in Section 4.

2. Building extraction

The first step extracts the 2D-supports of the urban structures from a DEM. Each 2D-support is represented by an unspecified quadrilateral (or triangle) and is associated with a specific part of a building. The 2D-supports of a building correspond to a set of connected quadrilaterals (i.e. non-overlapping quadrilaterals with common edges). We propose two different ways for the user to extract the 2D-supports: an interactive one and an automatic one.

Interactive extraction An operator controls interactive extraction: four clicks add a 2D-support (these clicks are the four points of the quadrilateral). This method allows accurate extraction as we can see on Figure 2, but takes considerable operator time.

Automatic extraction The second method is automatic: Ortner *et al.* obtain 2D-supports using an object approach based on marked point processes [10]. This method generates a set of rectangles which provide a rough description of the 2D-supports. Then, the rectangles are transformed into connected quadrilaterals by fusing neighboring rectangles. This process is automatic but the supports are not as accurate as those from interactive extraction - see Figure 2.

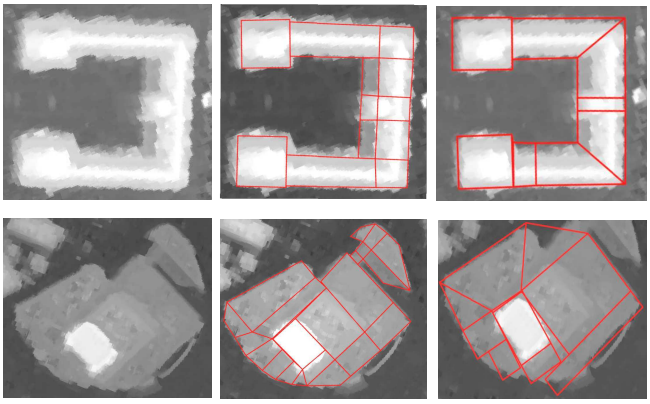


Figure 2. Extraction of 2D-supports: DEMs *(left)*, interactive *(center)*and automatic extraction *(right)*.

3. 3D reconstruction

Once the 2D-supports have been extracted, the buildings are automatically reconstructed through a density formulation. The first step specifies the 3D objects.

3.1. Library of 3D-blocks

The content of the library is a key point: if it is too limited (*e.g.* Lafarge *et al.*[7] had only flat and gable roof forms), the method loses generality. The proposed library, denoted by $\mathcal{M}$ and presented in Figure 3, allows us to reconstruct a large range of buildings through a collection of blocks. Each block possesses both a roof form and a variant:

- The proposed roof forms (denoted by $\mathcal{F}$ and illustrated in Figure 3*(top)*) include monoplane ($\mathcal{F}_{1x}$), multiplane ($\mathcal{F}_{2x}$) and curved roofs ($\mathcal{F}_{3x}$). Each roof form has a specific set of parameters F (the number of parameters varies between 1 and 6).

- The variants (denoted by $\mathcal{V}$ and shown in Figure 3*(bottom)* for a gable roof type) are specific to a roof form. They correspond to types of structure ends (hipped or straight ends) or structure junctions (”-”,”L”,”T” or ”+” junctions). The variants also specify the orientation of the roof with respect to the quadrilateral 2D-support. There are 1 to 3 parameters in the set of the variant, denoted by V.

To sum-up, each block of the library $\mathcal{M}$ is defined by a tuple $(\mathcal{F}, \mathcal{V})$ and an associated parameter set $\theta = (F, V)$. Some blocks can topologically be degenerated in some situations[2]: these cases are not allowed in the process in practice. Details concerning this library are available in [6].

[2]For example, a semi-elliptic roof on a triangular support

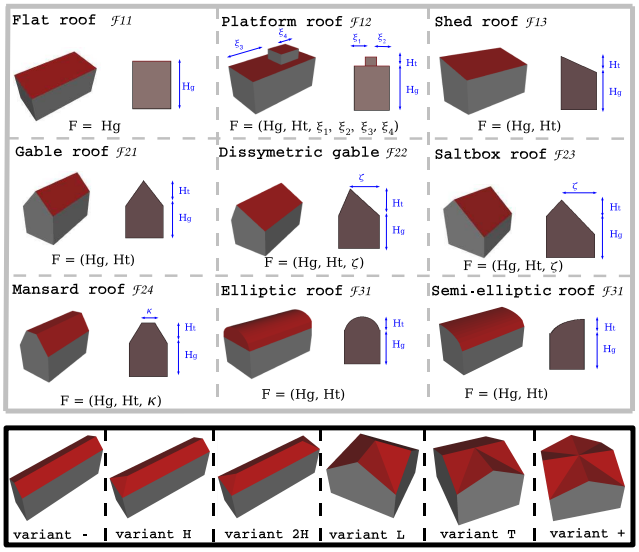


Figure 3. Library of 3D-blocks - the roof forms (3D and profile views) *(top)* and the variants of a gable roof *(bottom)*

3.2. Bayesian formulation

The notation for the Bayesian formulation is:
• S, a set of sites and $\Lambda = \{\Lambda(s)/s \in S\}$, a given DEM where $\Lambda(s)$ represents the intensity of the site, s.
• $\mathcal{C}$, the quadrilateral configuration representing the building 2D-supports associated with Λ. N is the number of quadrilaterals (see Figure 2).
• S_i, the subset of S whose sites are inside the quadrilateral, $i \in \mathcal{C}$.
• $\mathcal{D} = \{\Lambda(s)/s \in S_i\ , i \in \mathcal{C}\}$, the set of data.
• x, an element of the state space, $\mathcal{T}$, which corresponds to a configuration of 3D-parametric blocks knowing the 2D-supports $\mathcal{C}$. $x = (x_i)_{i\in\mathcal{C}} = (m_i, \theta_i)_{i\in\mathcal{C}}$ where each block, x_i, is specified by both a model, m_i of the library $\mathcal{M}$ and an associated set of parameters, θ_i. In the following, $x_i = (m_i, \theta_i)$ and m_i will be referred to as an object (or block) and a model, respectively.
• $\mathcal{S}_{x_i}$, the function from S_i to $\mathbb{R}$ which associates the roof altitude of the object, x_i, to each site of S_i.

We consider the random variable, X, distributed in $\mathcal{T}$, which follows an unnormalized density, h. h is the posterior density of a configuration, x, of objects, given $\mathcal{D}$. In a Bayesian framework, this density can be obtained from:

$$h(x) = h(x/\mathcal{D}) \propto h_p(x)\mathcal{L}(\mathcal{D}/x) \quad (1)$$

We must be able to build both a prior density, $h_p(x)$, and a likelihood, $\mathcal{L}(\mathcal{D}/x)$.

3.2.1 Likelihood

The likelihood represents the probability of observing the data, $\mathcal{D}$, knowing the configuration, x. By considering the hypothesis of conditional independence, it can be expressed in terms the local likelihood of objects, $\mathcal{L}(\mathcal{D}_i/x_i)$:

$$\mathcal{L}(\mathcal{D}/x) = \prod_{i\in\mathcal{C}} \mathcal{L}(\mathcal{D}_i/x_i) \propto \prod_{i\in\mathcal{C}} \exp -\Gamma^{\alpha}_{(i)}(\mathcal{S}_{x_i}, \mathcal{D}_i) \quad (2)$$

where $\Gamma^{\alpha}_{(i)}(.,.)$ is the distance from $\mathbb{R}^{card(S_i)} \times \mathbb{R}^{card(S_i)}$ to $\mathbb{R}$ defined by :

$$\Gamma^{\alpha}_{(i)}(\mathcal{S}_{x_i}, \mathcal{D}_i) = \left(\sum_{s\in S_i} |\mathcal{S}_{x_i}(s) - \Lambda(s)|^{\alpha} \right)^{\frac{1}{\alpha}} \quad (3)$$

To sum-up, the likelihood corresponds to the Z-error of the L_α norm between the DEM and the objects. In practice, $\alpha = \frac{3}{2}$ is a good compromise between sensitivity and robustness to DEM errors.

3.2.2 Prior

The prior density introduces interactions between neighboring objects. It allows us to both assemble objects in order to propose a realistic building and compensate for the lack of information contained in the DEM. A neighborhood relationship on $\mathcal{C}$ must be set up to define the interactions: two distinct quadrilaterals, i and $j \in \mathcal{C}$, are said to be neighbors if they have a common edge. The neighborhood relationship, denoted by $\bowtie$ ($i \bowtie j$ represents the set of neighboring pairs in $\mathcal{C}$). In previous work, too many interactions were set up [7]. The number must be minimal to preserve robustness and avoid problems in parameter setting. We propose a simple and efficient prior which is defined through a single interaction.
To do so, we define an assembly law which tests whether two objects can be assembled together. Two objects $x_i = (m_i, \theta_i)$ and $x_j = (m_j, \theta_j)$ are said "**joinable**" (denoted by $x_i \sim_a x_j$), if:

1 - $\mathcal{F}_i = \mathcal{F}_j$ and

2 - roof top orientations are compatible and

3 - the common edge of the quadrilateral 2D-supports, i and j, is not a roof height discontinuity.

The first condition checks that the two blocks have the same roof form. The second and third conditions test whether the roof tops of the two objects can be connected.
The prior favors "joinable" objects, *i.e.* homogeneous structures, which means that heterogeneous structures can also be reconstructed. Moreover, in order to avoid artefacts, the parameters of two "joinable" objects are encouraged to

have similar values. To do so, the unnormalized density, h_p, is expressed through a Gibbs energy, U_p (i.e. $h_p(x) = \exp -U_p(x)$), defined by:

$$\forall x \in \mathcal{T},\ U_p(x) = \beta \sum_{i \bowtie j} \mathbb{1}_{\{x_i \sim_a x_j\}} g(x_i, x_j) \qquad (4)$$

where $\mathbb{1}_{\{.\}}$ is the characteristic function. The parameter, $\beta \in \mathbb{R}^+$, weights the importance of the prior density with respect to the likelihood. The density only depends on this parameter which is computed using the Maximum Likelihood estimator under regularization constraints detailed in [6]. The function, g, taking values in $[-1, 0]$, measures the distance between the parameters of two "joinable" objects:

$$g(x_i, x_j) = \frac{D(x_i, x_j)}{D_{max}} - 1 = \frac{\sum_k \omega_k |\widetilde{\theta}_{i,(k)} - \widetilde{\theta}_{j,(k)}|}{D_{max}} - 1 \qquad (5)$$

$\widetilde{\theta}_{i,(k)}$ and $\widetilde{\theta}_{j,(k)}$ are the k^{th} element of the parameter sets F_i and F_j of the objects x_i and x_j, respectively. $D_{max} = \max_{x_i, x_j} D(x_i, x_j)$ is the maximum value of the distance. ω_k are weights which are introduced in this distance in order to normalize the parameter values according to the metric system. These weights are computed from the X, Y and Z resolutions and the configuration of quadrilaterals $\mathcal{C}$.
Figure 4 illustrates this interaction. If the two blocks belong to different roof types (for example a mansard roof and a semi-elliptic roof on the top right) or if the two objects do not have compatible roof orientations (bottom right), they will not be "joinable" and the energy will be null. On the contrary, if the two objects are "joinable", the energy will be negative : these configurations are favored. The nearer the parameters of the two objects, the lower the energy. The left configuration is the best one.

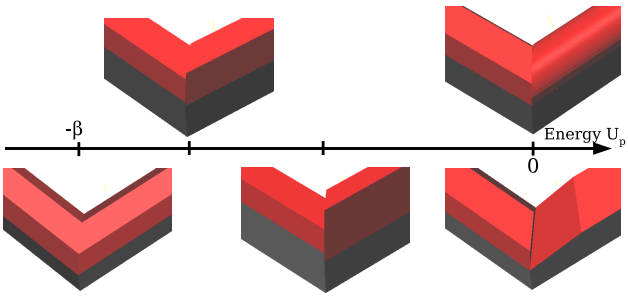


Figure 4. Prior energy - local configurations of various energies.

3.3. Optimization

We now find the object configuration which maximizes the posterior density,$h(.)$, i.e. the Maximum A Posteriori (MAP) estimator,x_{MAP}. This is a non convex optimization problem in a high and variable dimension space,$\mathcal{T}$, since the blocks in library, $\mathcal{M}$, are defined by a different numbers of parameters.

3.3.1 RJMCMC sampler

The Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm [4] is well adapted to our problem. Several papers have shown the efficiency of the RJMCMC sampler for the problem of multiple parametric object recognition. For example, Dick *et al.* use such a sampler to reconstruct architectural buildings from terrestrial images where the parametric object set includes structures such as columns, buttresses, entablatures or drainpipes [2].
The RJMCMC sampler simulates a discrete Markov Chain $(X_t)_{t \in \mathbb{N}}$ on $\mathcal{T}$ having π as an invariant measure (specified by the posterior density, $h(.)$) which makes small jumps between spaces of variable dimension respecting the reversibility assumption of the chain.
The jumps are proposed according to the three kernels specified in the following and detailed in [6]:

- **Kernel Q_1: uniform jumps** This is the classic kernel proposing a new state according to uniform distributions. It is enough to ensure that the Markov chain can visit any configuration of the state space. However, using only this kernel requires long computation time [7]. Thus we propose two efficient additional kernels Q_2 and Q_3.
- **Kernel Q_2: data-driven jumps** This kernel cleverly explores the state space using a data-driven process [13]. To do so, the state, x, is proposed knowing the data, i.e. according to a probability, $p(x|\mathcal{D})$. More precisely, it firstly estimates the gutter roof height $\widehat{H_g}$ and the roof top height $\widehat{H_t}$ of the object concerned by the jump and, secondly, chooses its height values according to the Gaussian distributions, $\mathcal{N}(\widehat{H_g}, \sigma)$ and $\mathcal{N}(\widehat{H_t}, \sigma)$, respectively (in practice, $\sigma = 1$ m).
- **Kernel Q_3: regularization jumps** In our application, the visual aspect of the result is very important: we need a kernel which proposes well-regularized objects, i.e. objects which are perfectly aligned with their neighbors. The new object, x_i, must be proposed knowing the neighboring objects $\{x_j / j \bowtie i\}$, i.e. according to $p(x_i|\{x_j / j \bowtie i\})$. The model is uniformly selected according to the models of the neighboring objects. The parameter values are chosen according to Gaussian mixtures depending on the parameter values of the neighboring objects.

In summary, the RJMCMC sampler is: at iteration t, if $X_t = x$:
1- Choose the kernel $Q_i(x, .)$ with probability q_i.
2- According to Q_i, propose a new state y.

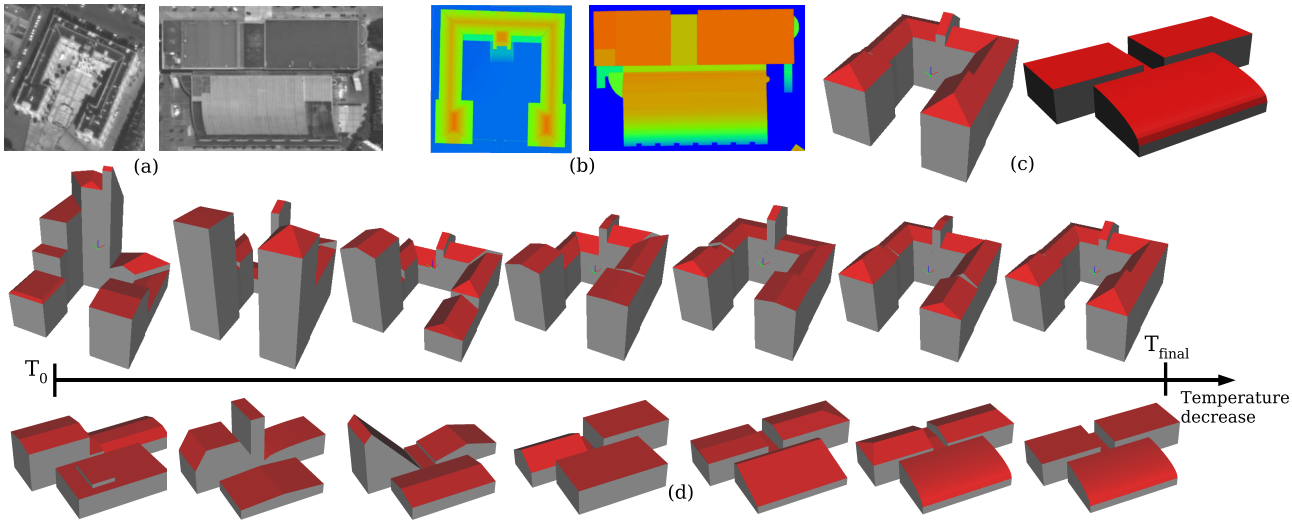


Figure 5. Two simple examples of the optimization process - evolution of the configuration as the temperature decreases *(d)*, associated with the satellite images *(a)*, the ground truths *(b)* and the final results *(c)*.

3- Take $x^{(t+1)} = y$ with probability:

$$\min\left(\frac{\pi(dy)}{\pi(dx)}\frac{Q_i(y,dx)}{Q_i(x,dy)}, 1\right) \tag{6}$$

and take $x^{(t+1)} = x$ otherwise.

3.3.2 Simulated annealing

Simulated annealing is used to ensure convergence: the density, $h(.)$, is substituted with $h(.)^{\frac{1}{T_t}}$, where T_t is a sequence of temperatures which tends to zero as t tends to infinity. Simulated annealing theoretically ensures convergence to the global optimum for any initial configuration, x_0, using a logarithmic temperature decrease. In practice, we used a geometric decrease which is faster and gives an approximate solution close to the optimal one. The initial and final temperatures are estimated through the variation of the energy following White [15].

The process has two stages. At the beginning, i.e. when the temperature is high (see Figure 5-(d) showing two simple examples of simulations), the process explores the density modes and favors configurations which have a high density. In this exploration stage, the data-driven kernel Q_2 is mainly used ($q_1 = q_3 = \frac{1}{8}q_2 = 0.1$)[3]. At low temperature[4], the configuration is close to the optimal solution and does not evolve very much : it involves a detailed adjustment of the 3D-block parameters. In this second stage, the regularization kernel Q_3 is mainly used ($q_1 = q_2 = \frac{1}{8}q_3 = 0.1$).

4. Experiments

The results were obtained from satellite DEMs (0.7 m resolution) and aerial DEMs (0.1 and 0.25 m resolution). The results show the reconstruction of complex buildings and dense urban areas whose level of detail depends on both the choice of the extraction process (automatic or interactive) and the kind of data. There is no comparison with other approaches for the reasons underlined in the introduction, except where the context is similar [7]. DEMs have been generated from 3-view images using a multi-resolution implementation of a Cox and Roy optimal flow matching image algorithm [11]. The 3D ground truths are raster images. Generic textures were applied to the objects of Figure 6 for visualization.

Figures 6 and 7 present various examples of reconstruction (showing different roof types, roof height discontinuities, closed structures or complex roof junctions). These results are convincing. The 3D-blocks are correctly assembled and few artefacts are generated which means the process adapts to buildings with complex roof junctions.

Results shown in Figure 6 were obtained from satellite DEMs with automatic 2D extraction. Even if some details are omitted, the shapes of buildings compare well to the ground truth and the generalization level is satisfactory with respect to the context. The different roof types were correctly identified and the roof height discontinuities were accurately located. The ground errors for the automatic 2D extraction process were satisfactory. The over-detection rate (in term of surface) was 9.7%. This rate can be improved by adding a vegetation mask to prevent the detection of trees.

[3] The q_i correspond to the probability of choosing the kernel Q_i in an iteration of the RJMCMC sampler.

[4] In practice, the second stage is detected when the accepted proposition rate computed on 1000 iterations becomes lower than 0.05.

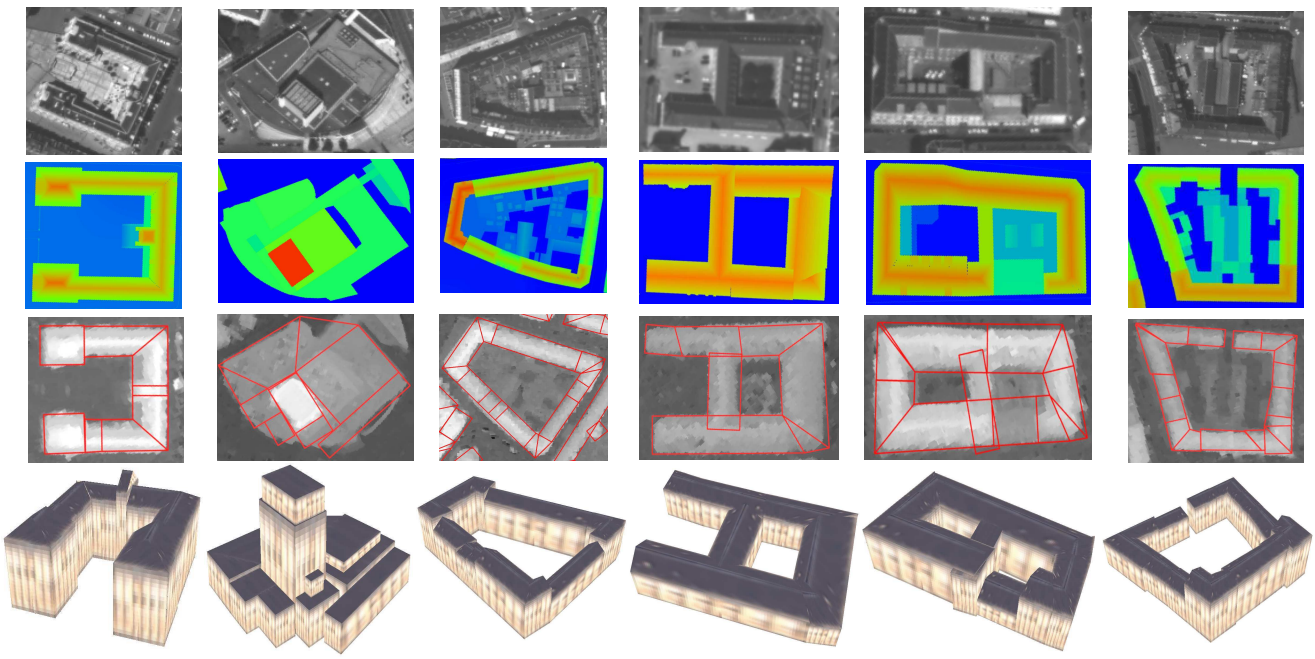


Figure 6. Reconstructed buildings with automatic 2D extraction from satellite data *(4^{th} row)*, satellite images *(1^{st} row)*, ground truths *(2^{nd} row)* and DEMs with 2D-supports *(3^{rd} row)*.

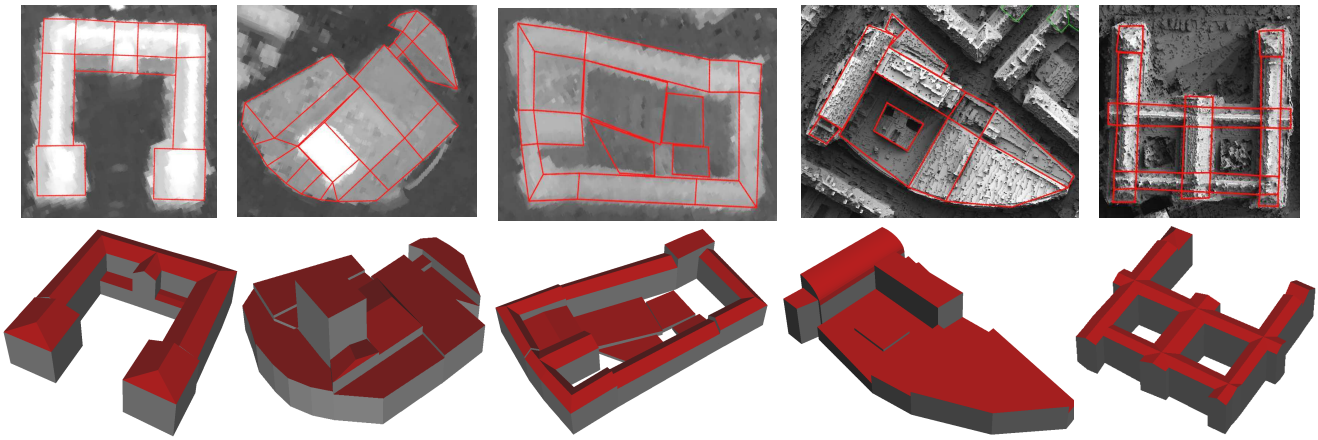


Figure 7. Reconstructed buildings with interactive 2D extraction from satellite data (3 first examples) and 0.25 m resolution aerial data (2 last examples) *(2^{nd} row)*, DEMS with 2D-supports *(1^{st} row)*.

The rate of missed detection was quite high (15%). However, it was mainly due to low flat buildings in inner courtyards (one floor height structures) that the 2D extraction process cannot detect since these buildings have low DEM discontinuities. Without taking into account these low flat buildings, this rate falls to 4.5%. The altimetric Root Mean Square Error (RMSE) in this context (satellite data / fully-automatic process) is 2.3 m. This is better than the 3.2 m error obtained earlier [7] in the same context, but still remains high. It is mainly due to both a non optimal positioning of 2D-supports in the automatic extraction process (which engenders important local altimetric errors at some locations) and inaccuracies in the DEMs (which mainly correspond to matching problems of non Lambertian surfaces such as glass rooves). The second example underlines the limits of the automatic 2D extraction process: some footprints (especially curved footprints) cannot be modeled accurately by sets of quadrilaterals. The proposed kernels achieved ac-

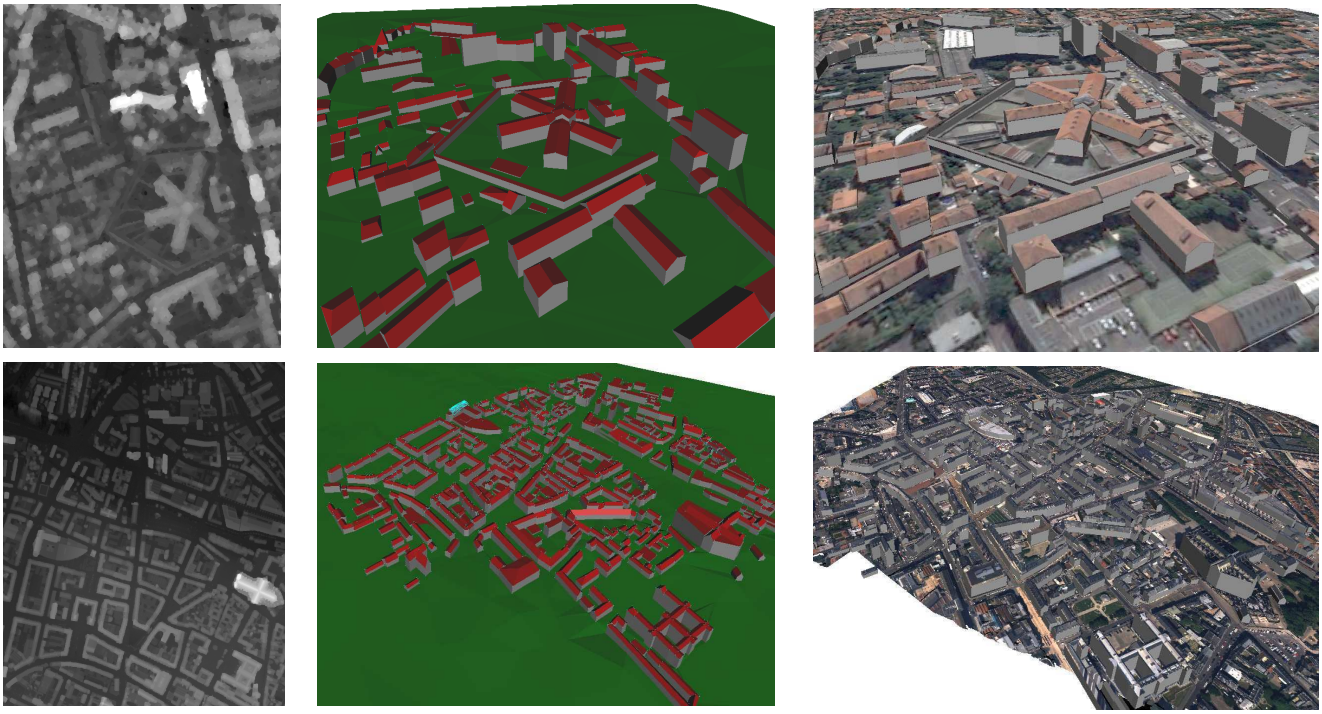


Figure 8. Urban area reconstruction without and with textures *(3^{rd}/2^{nd} columns)* from satellite data (1^{st} row) and aerial data (2^{nd} row), DEMs *(1^{st} column)*.

ceptable computation times. Less than one minute is necessary to obtain the buildings of Figure 6 using a 3Ghz processor (*vs* 5 minutes [7]).
Figure 7 presents examples of buildings reconstructed by the interactive 2D extraction process. In the two first examples, the results are clearly better since the 2D-supports are accurately located. The reconstruction is more detailed and the altimetric RMSE is 1.1 m from the satellite data.
Figure 8 shows results on two typical European down town areas. The computation time for the second example was 35 minutes ($0.6km^2$ - about 700 objects).
The proposed method also allows modeling roof details such as chimneys or dormer-windows. In fact, the library of 3D blocks is general enough to reconstruct such roof superstructures. Figure 9 presents accurate results of building superstructure reconstruction from a 0.1 m resolution aerial DEM. The altimetric RMSE for these images is 0.6 m - a very good value.

5. Conclusion

This new method is an interesting alternative to generic and parametric approaches. It presents several important characteristics. First, it obtains very good results from a single DEM. Moreover, contrary to other methods, this approach works efficiently on various data resolutions: a global description of the buildings from 0.7 m resolution satellite data or a detailled building reconstruction including roof superstructures from 0.1 m aerial data. The user can also choose the level of automation of the process since the 2D-supports can be extracted either interactively or automatically. To conclude, it is an adaptive method since other 3D-block types can be added to the library depending on the context.
In future work, it would be interesting to improve the optimization step to achieve both higher precision and shorter computing time. Adaptive cooling schedules could be used in the RJMCMC sampler or the Jump-Diffusion processes which are efficient for similar optimization problems. Moreover, we should evaluate the potential of this method on other kinds of cities such as typical North American urban areas.

References

[1] C. Baillard, C. Schmid, A. Zisserman, and A. Fitzgibbon. Automatic line matching and 3D reconstruction of buildings from multiple views. In *ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery*, 1999. 1

[2] A. Dick, P. Torr, and R. Cipolla. Modelling and interpretation of architecture from several images. *IJCV*, 60(2), 2004. 4

[3] C. Frueh, S. Jain, and A. Zakhor. Data processing algorithms for generating textured 3D building facade meshes from laser scans and camera images. *IJCV*, 61(2), 2005. 1

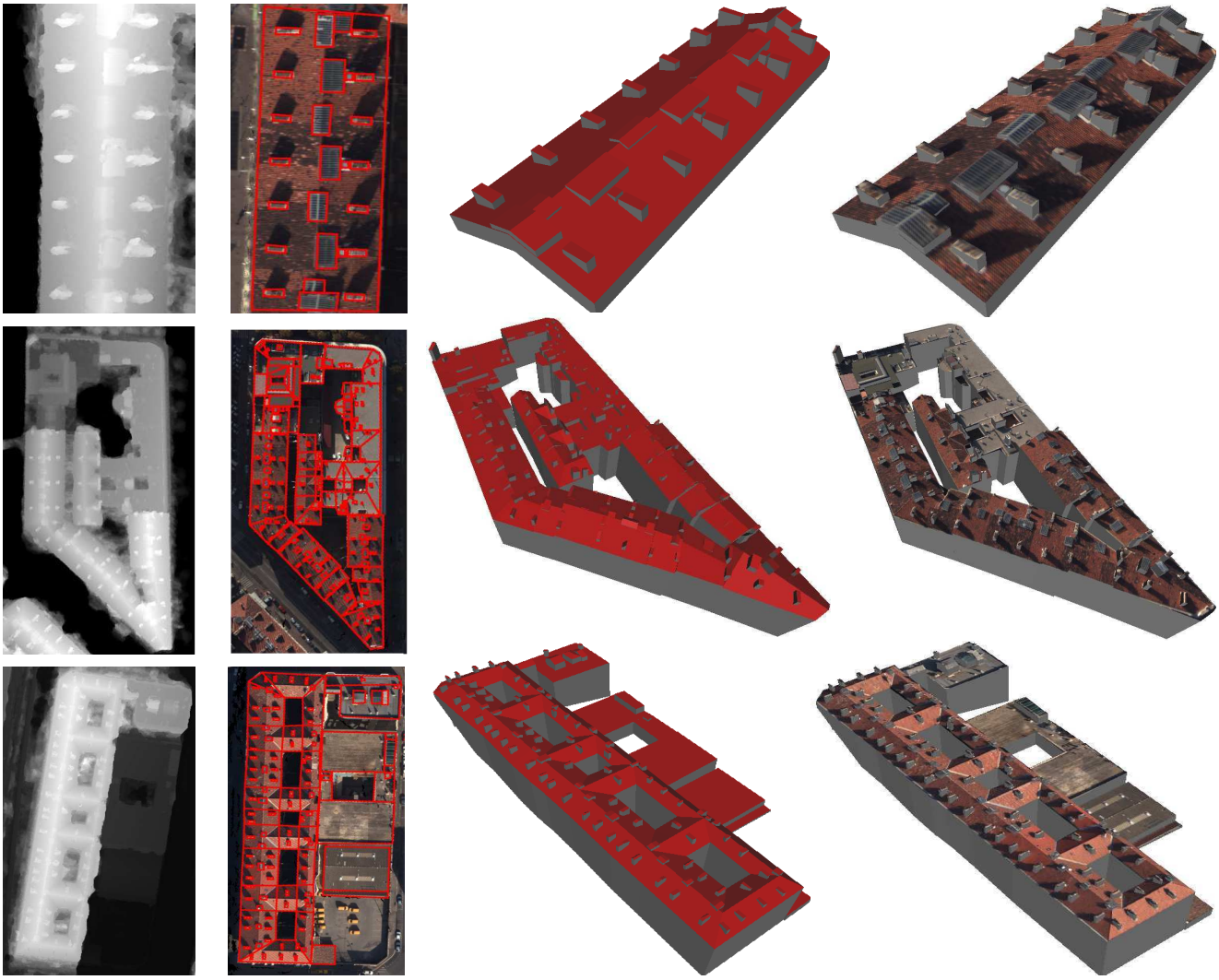


Figure 9. Building superstructure reconstruction without and with textures *(3^{rd}/4^{th} columns)* from 0.1 m resolution aerial DEMs *(1^{st} column)* and 2D-supports *(2^{nd} column)*.

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

[5] Z. Kim and R. Nevatia. Automatic description of complex buildings from multiple images. *CVIU*, 96(1), 2004. 1

[6] F. Lafarge. *Stochastic approaches for 3D building reconstruction*. PhD thesis, Ecole des Mines de Paris, Paris, France, 2007. 2, 4

[7] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny. An automatic building reconstruction method : A structural approach using high resolution images. In *Proc. IEEE ICIP*, Atlanta, USA, 2006. 2, 3, 4, 5, 6, 7

[8] S. Lee and R. Nevatia. Interactive 3D building modeling using a hierarchical representation. In *IEEE Workshop on higher-level knowledge in 3D modeling and motion*, Nice, France, 2003. 1

[9] P. Mueller, P. Wonka, S. Haegler, A. Ulmer, and L. Van Gool. Procedural modeling of buildings. In *Proceedings of ACM SIGGRAPH*, Boston, U.S., 2006. 1

[10] M. Ortner, X. Descombes, and J. Zerubia. Building outline extraction from Digital Elevation Models using marked point processes. *IJCV*, 72(2), 2007. 2

[11] S. Roy and I. Cox. A maximum-flow formulation of the n-camera stereo correspondence problem. In *Proc. IEEE ICCV*, Bombay, 1998. 5

[12] S. Scholze, T. Moons, and L. Van Gool. A probabilistic approach to building roof reconstruction using semantic labelling. In *Proceedings of the 24th DAGM Symposium on Pattern Recognition*, London, UK, 2002. 1

[13] Z. Tu and S. Zhu. Image Segmentation by Data-Driven Markov Chain Monte Carlo. *IEEE PAMI*, 24(5), 2002. 4

[14] U. Weidner and W. Forstner. Fowards Automatic Building Reconstruction from High Resolution Digital Elevation Models. *JPRS*, 50(4), 1995. 1

[15] S. White. Concepts of scale in simulated annealing. In *Proc. IEEE Int. Conf. on Computer Design*, 1984. 5