BUILDING EXTRACTION FROM DIGITAL ELEVATION MODELS

Mathias Ortner, Xavier Descombes, Josiane Zerubia

Ariana Research group
INRIA-CNRS-UNSA
2004 routes des lucioles,
06902 Sophia-Antipolis Cedex France BP 93
firstname.lastname@inria.fr

ABSTRACT

To extract buildings from Digital Elevation Models, we define a point process whose points represent buildings. We then define a density for this point process which is split into two parts, consisting in an "internal field" that allows us to model the prior knowledge we have on patterns of buildings in urban areas and an "external field" that makes the point process fit the data.
We then use a Metropolis Hastings Green sampler coupled with a simulated annealing that gives the configuration of buildings minimizing the energy we have defined.
We present results on real data provided by the French Mapping Institute (IGN).

1. INTRODUCTION

1.1. Building detection

Detecting buildings from aerial images and automatic reconstruction of urban scenes have become of deep interest in many applications : cartography, flight simulations, etc...

However, high density of urban areas and complexity of human made objects make it difficult to achieve, and automatic 3D urban area cartography is still an open problem. Some work has been done on this kind of task (see [2, 8, 9] for general overviews). Part of it focus on buildings and especially on roof modeling (see [2] to get a good overview of models and methods). Usually, these works are restrained to sparse areas, with just a few buildings, and the main goal is to describe complex buildings. The other part of the works focus on building detection.

Herein, we are interested by this second objective. Several methods and ideas have been proposed since a couple of years. A lot of them are based on primitive detection on two or more views. They rely mostly on line or corner detections and hypothesis testing.

Other works rely on DEM [1] construction using two or more views given by aerial images (see [3], for instance). The main idea is to detect buildings and to construct a DEM of the scene simultaneously. This is mixed with steps that use geometric and radiometric information to distinguish ground from buildings and buildings from vegetation. Some tentatives have been made by using only DEM (see [7] to have an example).

The first author would like to thank the French Defence Agency (DGA) and CNRS for partial financial support. The authors thank the French Mapping Institute (IGN) for providing the data.

[1] DEM : Digital Elevation Model.

1.2. Digital Elevation Models and Object representation

Our first goal is to refine DEMs. To achieve this, we have to use geometrical constraints. We want to obtain a vector representation of a dense urban area, for compression reasons, and because an object representation is close from a semantic one. We also want to be able to deal with low quality data. So we need to use a prior knowledge on urban areas in order to complete the missing information.

In [4], Garcin et al. use a point process approach to construct an object representation of an urban area. Point processes give an object-oriented approach allowing to deal with geometrical objects. Moreover it is possible to add interactions between points. Thus, it is possible to model the prior knowledge we have on the behavior of buildings in urban areas by viewing buildings as interacting particles within the point process framework. This has led us to use point processes to refine DEMs.

The data we use contains less information than radiometric images which are used in some other works on the same topic. However, DEMs differ only with respect to resolution and noise. By using DEM, we avoid problems seen while using radiometric information from images, since we do not need a lightening model of the scene.

Finally, it is worth to point out that refining DEMs is also of great interest in itself, since crude DEMs can be obtained not only by using stereo vision on radiometric images, but also by SAR interferometry or LASER measurement. In those cases, there is no radiometric information.

2. A MODEL FOR URBAN AREAS

We have seen that the marked point process framework fits our application well. Within this framework, an urban area is modeled by a set of an unknown number of points, each point standing for a building. We choose to model a building by a rectangular silhouette, a cost function telling how relevant is a proposed silhouette and a roof estimator that build a roof for a given relevant silhouette.

2.1. Silhouette model

We choose to model the silhouette of buildings by rectangles. Obviously, rectangles can be described by elements of a 5 dimensional space, using a center, a length, a width and an orientation. The space is denoted by S, made of a compact set where points live $K = [0, X_{max}] \times [0, Y_{max}]$ and the space of marks representing orientation, length and width of a rectangle. $S = K \times M = K \times [-\frac{\pi}{2}, \frac{\pi}{2}] \times [L_{min}, L_{max}] \times [l_{min}, l_{max}]$

More complex models could be used, like polygonal ones for instance.

2.2. Cost function

The cost function is a mapping from the space of rectangles to $\mathbb{R}$.

Assuming the Digital Elevation Model is defined on K we note h the mapping from points of K to $\mathbb{R}^+$ describing the DEM. For a point p, $h(p)$ is the height in meters given by the DEM.

Given a set $u = \{u_1, \ldots, u_n\}$ of points of K, we note $\bar{u}$ its mean, that is :

$$\bar{u} = \frac{\sum_{i=1}^{i=n} h(u_i)}{n}$$

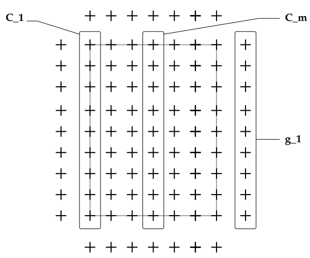


Fig. 1. Points used on a silhouette to compute the cost function and roof estimation.

Having a rectangle, we build a mask of points as shown on Figure 1. This mask of points is composed of 2 main areas. The first one is made of four lines around the rectangle $(g_1, \ldots, g_4)$ that are used to compute a ground height estimate $\hat{h}_g$. The second one is the central area c , which is made of N lines along the length $c_1, \ldots c_N$. In order to compute roof models, we make from these N lines m couples of lines $\{l_1, \ldots, l_m\}$:

$$m = \left(\frac{N+1}{2}\right), \quad \forall i \in \{1, \ldots, m\} \quad l_i = c_i \cup c_{N+1-i}$$

These couples of lines are symmetric with respect to the length-way axis of the rectangle and give us m means $m_i = \bar{l_i}$.

For the **ground estimate** $\widehat{h}_g$, we choose the lowest mean of the sets of points g_i. From the central area, we first define $v \in [0, 1]$ being the **volume rate** :

$$v = \frac{\text{card } \{p \in c \text{ s. t. } (h(p) - \widehat{h_g}) \geq h_{min}\}}{\text{card } c}$$

where h_{min} is a parameter of the model giving the minimal height we want to allow for a detected building. We also define the **homogeneity rate** $t \in [0, 1]$ being :

$$t = \frac{1}{m} \sum_{i=1}^{i=m} \frac{\text{card } \{p \in l_i \quad \text{s. t.} \quad |h(p) - m_i| < \sigma\}}{\text{card } l_i}$$

where σ is a parameter of the model related to homogeneity. We add a **surface rate** $s \in [0, 1]$ being :

$$s = \frac{l * L}{l_{max} * L_{max}}$$

Finally, we can define the cost function, obtained by trial and error, as a weighted product of these several rates. For a rectangle R and a DEM of height $h(.)$, we define the cost function J as :

$$J(R, h) = s * t^2 * v^3$$

This cost function has two intrinsic parameters : h_{min} which is a physical parameter, and σ that has to be tuned.

2.3. Roof estimation

There are quite a lot of possible roof models. The one we choose is simple and fits the cost function well : we model roofs with lines along the length, using the estimates of the cost function $m_0, \ldots, m_m$. This shape is symmetric, since the (m_i) stand for the mean estimates on symmetric lines along the length axis.

3. ENERGY MODEL

3.1. Marked point process

Working on S, we first consider a marked Poisson process on K and S, with intensity measure $(\lambda_K \times \mathbb{P}_M)(.)$ where $\lambda_K(.)$ is the Lebesgue measure on K and $\mathbb{P}_M(.)$ is a probability distribution on M. We note $\mu(.)$ the distribution of this Poisson process.

This process will be our reference process. We are going to define the distribution $\mathbb{P}_X(.)$ of our point process of interest X using the Radon-Nikodym derivative :

$$\frac{d\mathbb{P}_X}{d\mu}(\mathbf{x}) = \beta^{n(\mathbf{x})} f(\mathbf{x})$$

Here, β is a scale factor. The usual way of defining densities of point process is to write the density under its Gibbs form :

$$\mathbf{x} = \{u_1, \ldots, u_{n(\mathbf{x})}\} \quad u_i \in S \qquad f(\mathbf{x}) = \frac{1}{Z} e^{-U(\mathbf{x})}$$

where $U(\mathbf{x})$ is the energy of the set of particles $\mathbf{x}$. The lower the energy, the more probable the configuration. Therefore, the final goal of our framework consists in first building an energy, then finding configurations that minimize it.

3.2. Overlapping

We are dealing with points that represent the shape of a building. We know that buildings do not overlap. That is why we introduce in our model a **soft core** term, as used in [1] for instance.

Using a real parameter V_{inter} called **potential of intersection**, we define the energy linked to overlapping by $U_{inter}(\mathbf{x}) = V_{inter} * s(\mathbf{x})$, where $s(\mathbf{x})$ is an integer counting the number of pairs of rectangles that intersect in the configuration $\mathbf{x}$. If $V_{inter} > 0$, intersections are made repulsive since each new intersection increases the energy of the system.

3.3. Data term

The aim here is to use the cost function described before and the DEM in an energy term. This energy is given by $U_{data}(\mathbf{x}) = \sum_{u_i \in \mathbf{x}} V_d(u_i)$

The data term V_d should be quite smooth in order to ease the optimization. Moreover, minima of V_d should be relevant houses on the DEM. This is ensured by the construction of the cost function. Using two positive real numbers a, b, we propose the following function V_d living in $[-a, b]$:

$$V_d(R,h) = \begin{cases} b * (1 - \frac{v(R,h)}{v_{min}}^2) & \text{if } v(R,h) \leq v_{min} \\ -a * J(R,h) & \text{if } v(R,h) \geq v_{min} \end{cases}$$

for a given rectangle $R \in S$ and a height function $h(.)$. This data term has been obtained by trial and error and is quite robust. In practice, we use the following parameters (see section 5) :

$$\begin{array}{lll} a & = & 10 \\ b & = & 0.05 \end{array} \qquad \begin{array}{lll} v_{min} & = & 0.8 \\ h_{min} & = & 3 \text{ m} \\ \sigma & = & 1.5 \text{ m} \end{array}$$

For a low volume rate (ie. less than v_{min}) the rectangle is repulsive. But the closer v is from v_{min}, the less repulsive it is (smooth term in v^2). Then, when the volume rate is large enough, the homogeneity rate and the surface rate are involved in the data term. The surface rate makes bigger silhouettes be more valuable. The only important parameters are v_{min} since it gives the authorized error, h_{min} and σ. a and b allow to bound the energy of an object. This is important for mathematical reasons (see [11] for details.)

3.4. External Field, Internal Field and Temperature

The last step of the construction of the proposed model consists in using interactions and the data term in one energy. We call **internal field** the part of the energy related to the interactions (ie. the prior), **external field** the energy related to the data term, and **temperature** the coefficient T :

$$U(\mathbf{x}) = \frac{1}{T} \left(\rho . U_{data} + (1 - \rho) . U_{inter} \right), \quad \rho \in [0,1]$$

where ρ is the smoothing factor that tunes how important the data term is versus the internal field. This kind of weighting is very common in image processing.

4. OPTIMIZATION

Once we have defined the model, the next step consists in defining a procedure that allows us to find the configuration minimizing the energy. Here, this energy is related to a density of a point process, so the optimal configuration is the one that maximizes this density. This has been done several times, especially in image processing (see [1]). Usually, proposed algorithms are Monte-Carlo samplers coupled with simulated annealing. We follow this idea.

4.1. Monte Carlo Sampler for Point Process

The basic idea of such a sampler has been developed and justified by Geyer and Moller (see [5]), based on an extension of Hasting-Metropolis algorithm to point processes. The heart of the algorithm consists in birth or death of points.

However, this method produces a highly correlated Markov chain. As shown in the literature, while working with MCMC[2], one has to be careful with the proposition kernel involved in the sampler.

So, in order to improve the results, we use the work of Green in [6]. The idea is to add other transformations to the proposition kernel and to have a mixture of kernels. The 'birth or death' transformation has to remain, since it gives the mathematical properties we need.

The transformations we add to the birth and death, are translation, rotation and dilation. See [11] for mathematical details.

4.2. Simulated annealing

Once a sampler has been defined, it is possible to optimize the density, using simulated annealing. In the previous algorithm, we replace the temperature T by a time depending temperature T_t.

In theory, if we use the above sampler and make simultaneously decrease the temperature T_t from T_{init} to 0, (using a logarithmic law), the chain converges in total variation to a Dirac measure, whose mass is equally distributed on the global minima of $U(.)$.

Of course, since this kind of decrease is slow, we have to use in practice a geometric one and we loose the theoretical properties of global optimization.

5. RESULTS

We present below a result obtained on a part of a DEM of the city of Amiens (France). The resolution of this DEM is 20 cm by 20 cm horizontally and 15 cm vertically. Here, small intersections are allowed, for 3D visualization reasons. The results presented here have been obtained in 40 minutes with a `SUN-blade 2` (500 MHz, 250 MB). The image size is 1060 by 1024.

We can see that some small and low buildings are missing, because of the structure of the data which is too smooth when buildings are low. Figure 2 (c) shows the results of roof estimation and the ground truth (d) provided by the French Mapping Institute (IGN). This ground truth is precise but not complete since it was built by hand using high resolution (8 cm) aerial images.

As shown by Figure 2, some buildings are missing, and details of the shape of buildings are missed, due to the rectangular silhouette model we have chosen. Figure 3 shows the data and the result obtained by the proposed method in 3D.

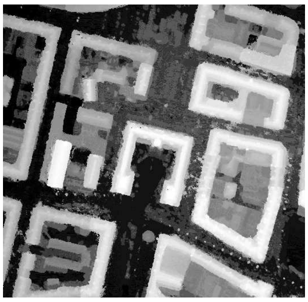
6. CONCLUSION AND FUTURE WORKS

The method we propose actually gives an automatic way of extracting buildings from a dense urban area by using a crude DEM as initial condition. The originality of this method can be appreciated as follows. First, the object oriented approach we have chosen gives an elegant way of adding geometric constraints during the extraction. Second, the way of adding a prior knowledge as interactions between buildings gives a nice framework to deal with poor quality data.

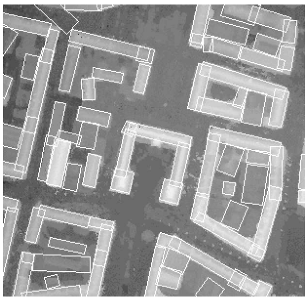
Of course, this work can be improved by testing the algorithm on more DEMs (optical, radar, laser etc...) and quantifying the quality of the result, improving the algorithm to be faster and to deal with larger areas, adding more complex models of buildings and roofs...

Tests have been done (see [10]) with more complex prior models, using relations such as alignment or orthogonality between buildings. Results are not improved with a more complex prior model due to the optimization step. Simulated annealing looses its efficiency if the energy exhibits too many local minima. Therefore, one way to improve the proposed technique is to work on a more efficient optimization algorithm. This will allow to refine

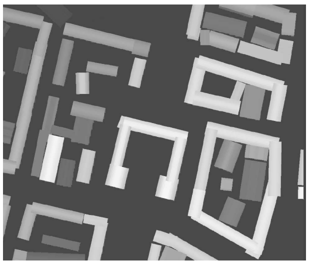
[2]MCMC : Monte Carlo Markov Chain.



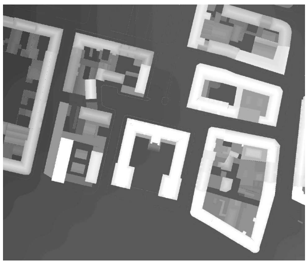
(a) crude DEM (IGN)



(b) building extraction.



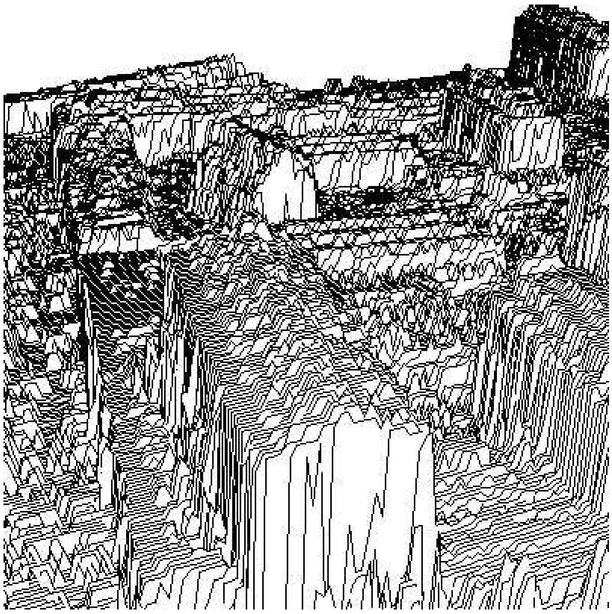
(c) estimated roofs



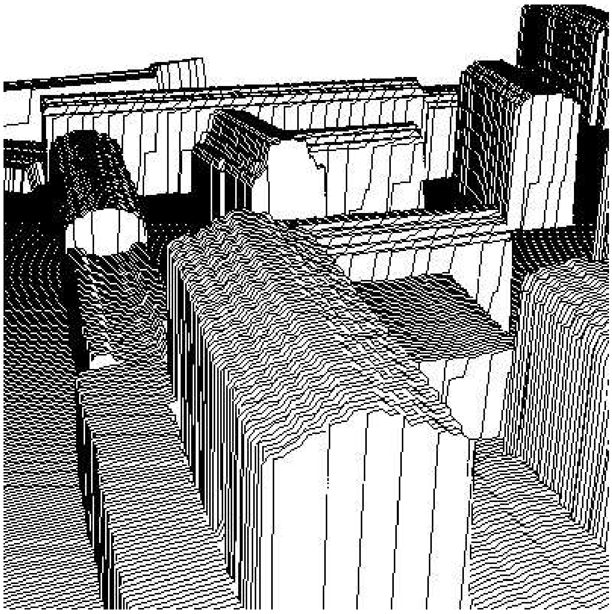
(d) ground truth (IGN)

Fig. 2. Experimental results.

the model and use more complex silhouettes and a bigger set of possible roofs. This should be done in a near future.



(a) crude DEM (IGN)



(b) refined DEM obtained by the proposed method

Fig. 3. Result in 3D

7. REFERENCES

[1] A. Baddeley and M. N. M. van Lieshout. Stochastic geometry models in high-level vision. In K.V. Mardia, editor, *Statistics and Images*, volume 1, pages 233–258. Abingdon: Carfax, 1993.

[2] A. Fischer, T. H. Kolbe, F. Lang, A. B. Cremers, W. Förstner, L. Plümer, and V. Steinhage. Extracting buildings from aerial images using hierarchical aggregation in 2D and 3D. *Computer Vision and Image Understanding: CVIU*, 72(2):185–203, 1998.

[3] M. Fradkin, M. Roux, and H. Maître. Building detection from multiple views. In *ISPRS Conference on Automatic Extraction of GIS Objects form Digital Imagery*, 1999.

[4] L. Garcin, X. Descombes, J. Zerubia, and H. Le Men. Building detection by Markov object processes and a MCMC algorithm. *INRIA Research Report 4206*, June 2001.

[5] C.J. Geyer and J. Møller. Simulation and likehood inference for spatial point process. *Scandinavian Journal of Statistics*, Series B, 21:359–373, 1994.

[6] P.J. Green. Reversible jump Markov chain Monte-Carlo computation and Bayesian model determination. *Biometrika*, 57:97–109, 1995.

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

[8] H. Mayer. Automatic object extraction from aerial imagery- a survey focusing on buildings. *Computer Vision and Image Understanding (CVIU)*, 74(2):138–149, 1999.

[9] R. Nevatia and K. Price. Automatic and interactive modeling of buildings in urban environments from aerial images. In *ICIP*, September 2002.

[10] M. Ortner. Extraction de caricatures de bâtiments sur des modèles numériques d'élévation. Master Thesis (DEA, in French), August 2001.

[11] M. Ortner, X. Descombes, and J. Zerubia. Building detection from digital elevation models. *INRIA Research Report 4517*, July 2002.