A DISCONTINUITY DETECTOR FOR BUILDING EXTRACTION FROM DIGITAL ELEVATION MODELS BY STOCHASTIC GEOMETRY

Mathias Ortner, Xavier Descombes, and Josiane Zerubia

Ariana - Joint research group CNRS / INRIA / UNSA
INRIA, 2004 route des Lucioles
BP 93, Sophia Antipolis, France
Email : Firstname.Lastname@sophia.inria.fr
Web : www-sop.inria.fr/ariana/

ABSTRACT

This work presents a framework to use stochastic geometry for automatic building extraction from different kinds of Digital Elevation Models (DEM). The goal is to extract some vectorial information from a DEM in order to ease precise 3 dimensional reconstruction. Using a spatial point process framework, we model cities as configurations of unknown number of rectangles. An energy is defined, which takes into account both a low level information provided by the altimetry of the scene, and some geometric knowledge on the location of buildings in towns. The estimation is done by minimizing an energy using simulated annealing. We present results on real data provided by IGN (French National Geographic Institute) consisting of laser and optical DEMs.

1. INTRODUCTION

For an increasing number of applications among which cartography, urban planning and military intelligence, automatic reconstruction of 3 dimensional precise maps of towns is an important but still open issue.

Several methods and algorithms have already been proposed to deal with urban areas using High Resolution (HR) remote sensing data. General overviews can be found in [2, 7] or in the introduction of [6]. Concentrating on automatic methods, a general synopsis of the existing approaches emerges. Automatic methods are mostly made up of three steps : focusing on an area of interest, low level primitive extraction and building reconstruction. Primitive detection and building reconstruction are usually deeply linked, since they hang on two complementary processes: the former creates aggregations of primitives (bottom-up step, as illustrated in [2]), while the latter matches building models from a data base with the previously obtained hypothesis (top-down step). During the matching step, these methods face a combinatorial problem. To ease this step, methods proposed in the literature tend to increase the number of data used such as multiple color images, Laser data, register maps or hyperspectral images (see examples in [3, 6]).

We present here a different approach and propose: to fuse the two first steps of focusing and primitive detection in an automatic method that extracts simple shapes of buildings, to only use the data given by a Digital Elevation Model (DEM) and to introduce some spatial knowledge on the location of buildings in dense urban areas. We do not deal here with the precise 3 dimensional reconstruction but with the extraction of the outlines of buildings. The same goal has been considered by S. Vinson in [13]. Our approach relies on a point process model. Such models have been introduced in image processing by A. Baddeley and M.N.M. van Lieshout in [1], and were later developped to tackle applied problems in image processing (for instance by H. Rue and M. Hurn in [11]).

2. MODEL DEFINITION

2.1 Spatial point process

Let consider a point process X living in $K = [0, X_{max}] \times [0, Y_{max}]$ which describes an image. X is a measurable mapping from a probability space $(\Omega, \mathscr{A}, \mathbb{P})$ to **configurations of points** of K:

$$\forall \omega \in \Omega \;\; X(\omega) = \{x_1, \dots, x_n\} \;\; x_i \in K$$

Basically, a point process is a random variable whose realizations are random configurations of points. K is actually considered as a torus.

Point processes were introduced in image processing because they easily allow to model scenes of objects. A marked point process adds some marks (parameters) to each point. Taking $S = K \times M$ with the following M allows to describe random configuration rectangles:

$$M = [-\frac{\pi}{2}, \frac{\pi}{2}] \times [L_{min}, L_{max}] \times [l_{min}, l_{max}]$$

We note $\mathscr{C}$ the set of all finite configurations of points of S.

Let consider the distribution $\mu(.)$ of a given Poisson point process which we call reference Poisson point process. Under some conditions (detailed in [4]) it is possible to define a point process X by specifying its probability distribution by a density with respect to the dominating distribution $\mu(.)$. Let consider a mapping $h(.)$ from the space of configuration of points $\mathscr{C}$ to $\mathbb{R}$, and a real parameter Z such that:

$$Z = \int_{\mathscr{C}} h(\mathbf{x}) d\mu(\mathbf{x})$$

If a point process X defined by an unnormalized density $h(.)$, and a reference intensity measure $\nu(.)$ (defining a reference Poisson point process distribution) is given, it is possible to build a Markov Chain that converges ergodically to the distribution of X. Once such a sampler is defined, any Monte Carlo value is computable. It includes mean values, variances etc... Another way of using this sampler is to use it within a simulated annealing framework which gives a global maximum of the density $h(.)$ as detailed in [12]. Thus, the corresponding estimator is given by:

$$\hat{\mathbf{x}} = Argmax \;\; h(.)$$

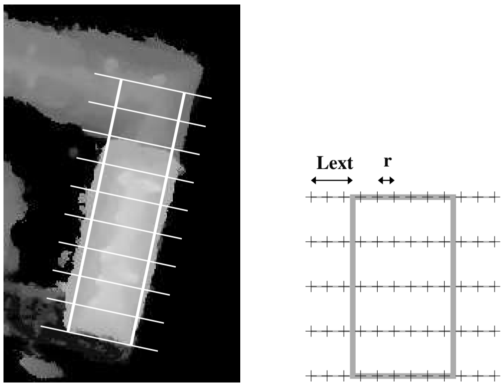
2.2 Point process of rectangles and buildings

We use random configurations of rectangles to describe towns. Rectangles seemed to be a natural pattern to be detected in dense urban areas. In [9] we presented a model based on two hypothesis on the homogeneity of heights inside a rectangle. Here we present a model based on discontinuity detection. This model relies on a low level filter which can be adapted to different kinds of DEMs.

2.2.1 General form of the energy

Writing the unnormalized density $h(.)$ under its energetic form $U(\mathbf{x}) = -log(h(.))$, we define a model that can be decomposed as:

$$\begin{aligned} U(\mathbf{x}) &= \rho U_{ext}(\mathbf{x}) + U_{int}(\mathbf{x}) + U_{excl}(\mathbf{x}) \quad (1) \\ &= \rho \sum_{u \in \mathbf{x}} U_d(u) + U_{int}(\mathbf{x}) + U_{excl}(\mathbf{x}) \quad (2) \end{aligned}$$



(a) Disposition of measured slices

(b) Sub-sampling

Figure 1: A DEM, a rectangle and slices construction

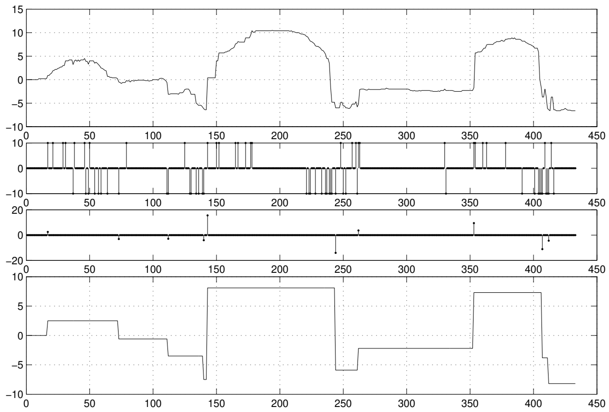


Figure 2: Steps of the profile simplification algorithm: From top to bottom : Profile, detected slopes ($\geq \sigma_l$), detected gradients after slope accumulation and opening steps and simplified profile obtained on the example of Amiens (IGN).

where $U_{int}(\mathbf{x})$ stands for an internal energy giving a spatial structure to the configuration $\mathbf{x}$, while $U_{ext}(\mathbf{x})$ is the external field quantifying the quality of a configuration given the data that can be decomposed as a sum of energies per object $U_d(.)$, and U_{excl} is an exclusion term to avoid superposing objects and is tuned such that :

$$U(\mathbf{x} \cup u \cup u) > U(\mathbf{x} \cup u) \quad \forall\, (u, \mathbf{x}) \in S \times \mathscr{C}$$

2.3 Data term

The purposes underlying the definition of the data term $U_d(.)$, mapping from S to $\mathbb{R}$, are first to decide what an "attractive object" ($U_d(u) \leq 0$) is, and second, to introduce a potential such that minimizing $U_d(u)$ locally gives the closest attractive object.

2.3.1 Low level filter

In an altimetric data, an important information comes from discontinuities. Fig. 1(a) shows a part of a DEM, a rectangle on it and slices evenly disposed. This slice gives some points of interest which are discontinuities. We propose to extract some points of interest from each slice and to check the coherence between these points and the rectangular shape of the object. Thus, to compute $U_d(u)$, we propose to use a low level algorithm that relies on the following idea. For a given rectangle $u \in S$, we propose to measure first all related slices and then, for each slice, to detect points of interest. We use three parameters e, r and L_{ext} that define a mask of points (shown on Fig. 1(b)) used to compute the profiles associated to a rectangle on the DEM. Then, to detect points of interest, the following procedure is applied to each profile. First, for each profile, slopes higher than a threshold σ_l / r are accumulated using a regularizing parameter l_{regul}. Then an opening is performed using the same structuring element (segment of l_{regul} length) . Finally, gradients larger than a threshold σ_h are selected, and thus, for each profile, a set of points of interest is drawn up. This procedure can be implemented in a fast way. The 4 parameters: r,σ_l, σ_h and l_{regul}, all defined in meters, are easily learned if a training set of profiles is provided. An important issue is that it is easy to relearn them when changing the kind of data used (for instance, when switching from a optical DEM to a Laser one.)

2.3.2 Values of interest

For a given rectangle, of first importance is the *hit length* value: it corresponds to the number of detected gradients which are close enough to the corresponding rectangle side multiplied by the resolution parameter e. Fig. 3(d) illustrates the meaning of the hit length $Lg(u)$ on an example: multiplying the number of boxed gradients by e (length of a box) gives the hit length, ie. the length of detected discontinuities on the DEM. $Lg(u)$ is computed by the mean of a sensibility parameter δr (width of a box). Since there are two sides, we actually compute two values that can be ordered: $L_{min} \leq Lg_1(u) \leq Lg_2(u) \leq L_{max}$. A *volume rate* $\bar{v}(u)$ and a *moment rate* $\bar{m}(u)$ are also computed. The volume rate corresponds to the rate of grey lines of Fig 3(b), while the moment rate corresponds to the rate of squared length of grey lines in Fig 3(c). They intuitively can be seen as a translation and a rotation potential measuring the distance between a rectangle hypothesis and the closest relevant building.

2.3.3 Energy definition

We define γ_0 the set of attractive objects as the set of objects with enough discontinuities detected along their sides. Let $th_1 \leq th_2$ be two thresholds, both living in $]0,1[$, $L(u)$ and $l(u)$ the length and the width of the rectangle u; γ_0 is then the following set :

$$\gamma_0 = \left\{ u \in S \quad \text{s.t.} \quad \begin{array}{c} Lg_1(u) \geq th_1 * L(u) \\ \text{and} \\ Lg_2(u) \geq th_2 * L(u) \end{array} \right\} \tag{3}$$

We associate to this set the energy function $J_0 : S \to [0, -1]$

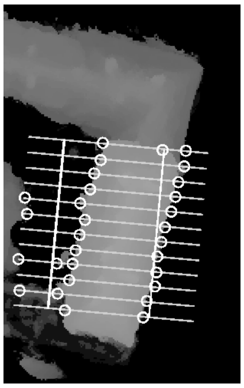
$$J_0 : \quad \begin{array}{lll} u \in \gamma_0 & \to & -\left(\frac{1}{4} \frac{Lg_1(u)+Lg_2(u)}{L_{max}} + \frac{1}{2} \frac{l(u)}{l_{max}} \right) \\ u \notin \gamma_0 & \to & \varepsilon(u) \end{array} \tag{4}$$

$\varepsilon(u)$ is strictly positive and actually depends on the volume and moment rates. It reflects a kind of distance between u and γ_0 in order to favor among repulsive objects those that are closer from γ_0. Details are provided in [8]. An object is thus attractive if and only if the detected hit lengths are large enough. Basically, we define $U_d(u) \approx J_0(u)$. Minimizing the cost function maximizes the length of detected gradients and the perimeter of the rectangle. Over the total configuration, since $U_d(\mathbf{x}) = \sum U_d(u)$, the reward function evolves linearly with the total detected discontinuity length.

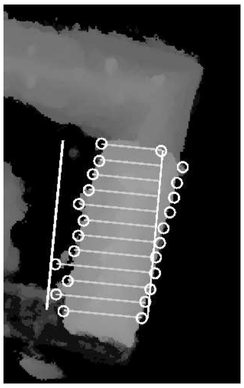
2.4 Internal Field

2.4.1 Attractive Interactions

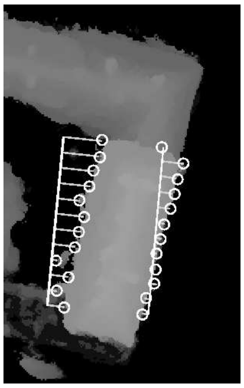
The prior model can be seen as a regularizing term. In our framework, the prior model is essentially composed of interactions between objects. The implemented internal field favors alignment between detected structures, completion and some paving behavior. Fig. 4 presents these three interactions. They all are defined



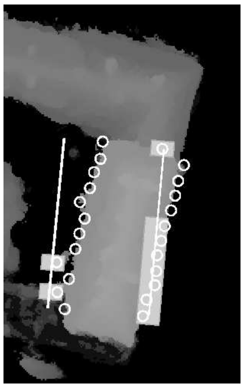
(a) **Detection :** Gradients detected by the profile simplification procedure.



(b) **Volume rate :** Grey lines represent segments used to compute the volume rate.



(c) **Moment rate :** Grey lines represent segments used to compute the moment rate.



(d) **Localization :** A gradient is boxed if it is close enough to the corresponding rectangle side.

Figure 3: A DEM, a rectangle hypothesis and the values computed to measure the relevance of this current rectangle hypothesis.

through conditions on respective angles of two close rectangles and the distance between appropriate corners.

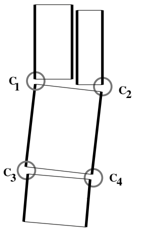
2.4.2 *Exclusion Interaction*

An exclusion term that avoids redundant objects is needed, because of the following reasons: we need to avoid redundant explanations of the data and we need to insure that the attractive interactions will not make the set of particles collapse to an infinite accumulation of points. Furthermore, a condition used in the convergence of the algorithm proof requires the variation of the energy, induced by adding a point to a given configuration, to be bounded.

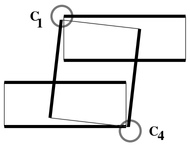
The simple exclusion interaction we use is the *intersection relation* that acts only on parallel rectangles.

2.4.3 *Visual results*

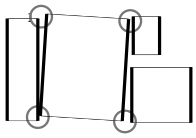
We present on Fig. 5 two results. The first one (Fig. 5(a)) is a realization of a point process whose density is defined by $(U_{int}(.) + U_{excl}(.))/T$ at a medium temperature. The second one (Fig. 5(b)) is a realisation of the same process at low temperature obtained by



(a) Alignment

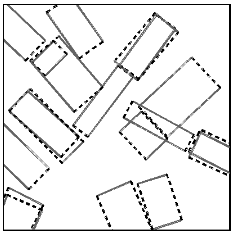


(b) Completion

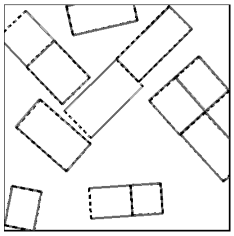


(c) Paving

Figure 4: Attractive interactions.



(a) At medium temperature



(b) At low temperature

Figure 5: Illustration of the influence of the internal field.

simulated annealing. These results show how the internal field constraints object positioning.

3. THE ALGORITHM

A simulated annealing is performed on the density of the defined point process, under its Gibbs representation. This simulated annealing requires a RJMCMC (Reversible Jump Markov Chain Monte Carlo) sampler, which has been provided by C. Geyer and J. Møller in [4], P.J. Green in [5], and Van Lieshout in [12]. We also worked on it in [10]. We use this work to build an efficient sampler that uses an inhomogeneous reference Poisson point process and birth or death in neighborhoods. The inhomogeneous part is useful since the ratio between $|\gamma_0|$ and $|S|$ is really small. We introduce three other levels partitioning the complement of γ_0, and tune the exploration ability of the Markov Chain with respect to these levels through the inhomogeneity.

4. RESULTS

We present here only two results. More results are available in [8]. Both results were obtained on real data and are presented on Fig. 6. The first result corresponds to a part of an optical DEM of Amiens. The size of the image is around 1050 by 1050. Resolutions are 20cm (X,Y) and 10 cm (Z). Computations took around 3 hours on a 1Ghz, 1GB Linux computer for 10 billions iterations. Overdetected parts represent 10.6%, while missed parts 8.9% of the total image. The second result was obtained from a LASER DEM of the same area. Size of the image is smaller, since the resolution is 50cm in XY directions. The overdetected surface represents 11% of the total area and the missed one 5%. The computational time was of the same order as the time needed for the optical DEM. It tends to prove that with such object models the scene complexity rather than the image size is important for the required number of iterations.

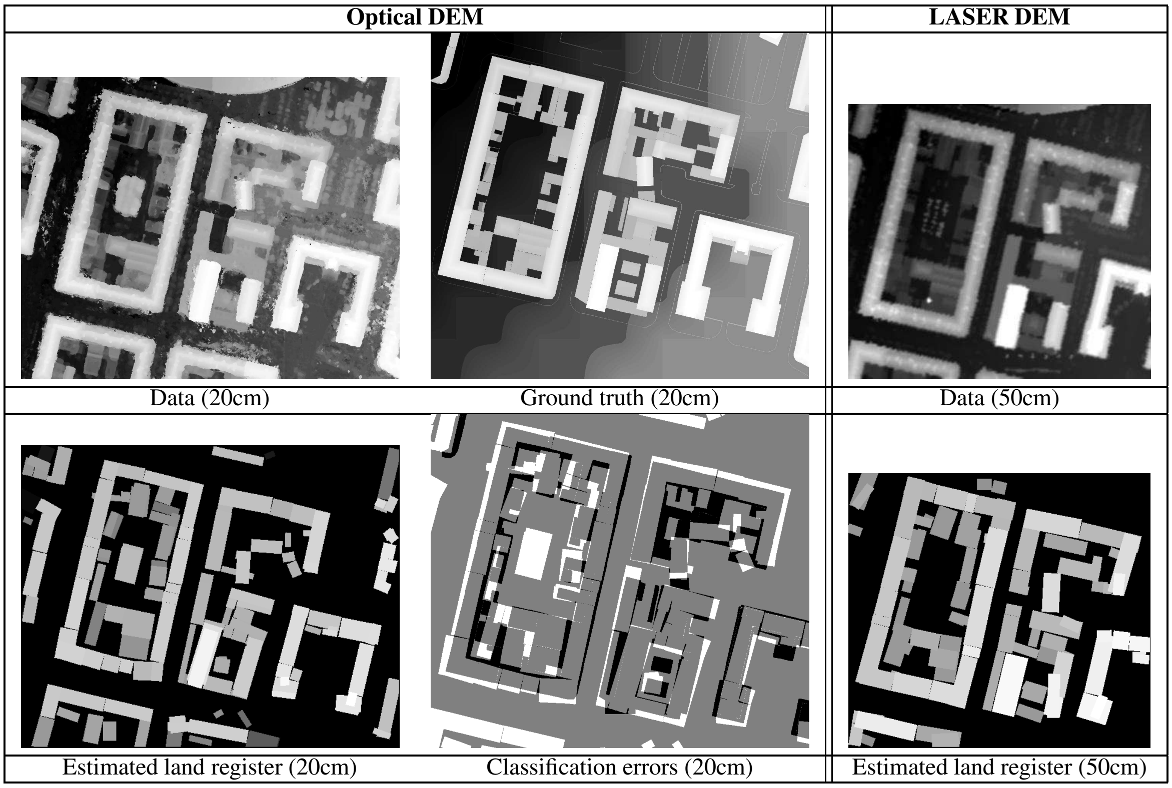


Figure 6: Estimated land registers on parts of an optical DEM (left) and a LASER DEM (right).

5. CONCLUSION

We have proposed an algorithm that turned out to be efficient for the difficult problem of extracting buildings from urban areas. The algorithm extracts simple rectangular shapes from altimetric data. It is an automatic approach which means that there is no need for any interaction with an operator and, in particular, no initial conditions are needed. The algorithm seems to be efficient even on dense urban areas, and the parametrization quite robust, since the same model is used for both LASER and optical DEMs.

Future work will involve implementing a MCMC-ML procedure to see if it is relevant for parameter estimation. We also plan to try to do data fusion and use more data. Another direction to be explored is the possibility of using more complex models of buildings to improve the quality of the obtained result.

6. ACKNOWLEDGMENTS

The authors would like to thank the the French National Geographic Institute (IGN) for providing the images of Amiens and the ground truth. The first author thanks the French Defense Agency (DGA) and CNRS for partial financial support.

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

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

[6] H. Jibrini. *Reconstruction automatique des bâtiments en modèles polyédriques 3D à partir de données cadastrales vectorisées 2D et d'un couple d'images aériennes à hautes résolution*. PhD thesis, ENST, Paris, France, 2002.

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

[8] M. Ortner, X. Descombes, and J. Zerubia. Automatic 3D land register extraction from altimetric data in dense urban areas. *INRIA Research Report 4919*, August 2003.

[9] M. Ortner, X. Descombes, and J. Zerubia. Building detection from digital elevation models. In *ICASSP*, volume III, Hong Kong, April 2003.

[10] M. Ortner, X. Descombes, and J. Zerubia. Improved RJMCMC point process sampler for object detection by simulated annealing. *INRIA Research Report 4900*, August 2003.

[11] H. Rue and M. Hurn. Bayesian object identification. *Biometrika*, 3:649–660, 1999.

[12] M. N. M. van Lieshout. Stochastic annealing for nearest-neighbour point processes with application to object recognition. *CWI Research Report, BS-R9306, ISSN 0924-0659*, 1993.

[13] S. Vinson, L. D. Cohen, and F. Perlant. Extraction of rectangular buildings using DEM and orthoimage. In *SCIA*, Bergen, Norway, June 2001.