# HAL
archives-ouvertes.fr

# A marked point process of rectangles and segments for automatic analysis of Digital Elevation Models

Mathias Ortner, Xavier Descombes, Josiane Zerubia

## ▶ To cite this version:

## HAL Id: inria-00278882
## https://hal.inria.fr/inria-00278882

Submitted on 14 May 2008

# A marked point process of rectangles and segments for automatic analysis of Digital Elevation Models.

Mathias Ortner, Xavier Descombes and Josiane Zerubia

*Abstract*—This work presents a framework for automatic feature extraction from images using stochastic geometry. Features in images are modeled as realizations of a spatial point process of geometrical shapes. This framework allows the incorporation of a priori knowledge on the spatial repartition of features. More specifically, we present a model based on the superposition of a process of segments and a process of rectangles. The former is dedicated to the detection of linear networks of discontinuities, while the latter aims at segmenting homogeneous areas. An energy is defined, favoring connections of segments, alignments of rectangles, as well as a relevant interaction between both types of objects. The estimation is performed by minimizing the energy using a simulated annealing algorithm.

The proposed model is applied to the analysis of Digital Elevation Models (DEMs). These images are raster data representing the altimetry of a dense urban area. We present results on real data provided by the IGN (French National Geographic Institute) consisting in low quality DEMs of various types.

*Index Terms*—Image processing, Poisson point process, stochastic geometry, dense urban area, Digital Elevation Models, land register, building detection, MCMC, RJMCMC, simulated annealing.

## I. Introduction

### A. Dense urban areas and building reconstruction

As cities are the place of increasing concentrations of people they are the centers of numerous interests: economical, military, environmental, to name a few. 3D cities representations are of first interest for different communities (telecommunication, security, etc...). However, automatically obtaining such representations is still an open issue.

The remote sensing community provides various sensors and techniques to accumulate data on a specific urban area. In particular, the advent of high resolution data (HR) in remote sensing has given aerial and satellite images a primary role in analyzing urban areas. Other imaging techniques like LASER or LIDAR sensors provide different data. However, the complexity of urban areas make the data challenging for automatic analysis.

In this work, we focus on the analysis of Digital Elevation Models (DEMs). DEMs are raster data representing the altimetry of an urban area. They consist of 2D images such that the gray level of a pixel describes the height of the corresponding point in the scene. This type of data is obtained through different sensors and processing techniques (e.g. stereovision,

LASER sensing, etc...). The sample presented in Figure 11 is a typical example of a Digital Elevation Model obtained by aerial stereovision.

Extraction of buildings from urban data has been subject to a corpus of literature. General overviews can be found in [1], [2] or in the introduction of [3]. Automatic methods are mostly made up of three steps. The *focalization step* selects a relevant area that supposedly corresponds to a building. A pre-segmentation (ground/above ground) or the incorporation of external vectorial data (e.g. a land register) are usual ways of achieving this pre-selection. The two next steps, namely the *primitive detection* and the *building reconstruction* are usually closely linked. The *bottom up* process creates aggregations of primitives (illustrated for instance, in [1]), while the *top-down* procedure matches the obtained aggregation hypothesis with building models. The association of hypothetical aggregations with pre-defined building models is a combinatorial problem and thereby concentrates most of the computational load of the procedure.

As the structure of dense urban areas is tremendously complex, there is a huge need for incorporating as much information as possible. The fusion of different types of data is a first possibility. Many proposed methods tend to increase the variety of data used by including multiple color images, LASER clouds of points, register maps or hyperspectral images (see examples in [3]–[5]). Additional information can also be integrated as *a priori knowledge*. Proposing a set of possible building shapes is a first way of incorporating prior knowledge. Our approach allows going further. We propose to incorporate knowledge on the patterns of the primitives to be extracted in terms of interactions between objects.

### B. Our approach: spatial point process models

Our approach consists in modeling an urban area by a set of an *unknown number* of interacting particles, where each particle stands for a building element. A particle is eventually a geometrical object that can be compared to the data.

In [6] we presented an original approach and model cities as realizations of a spatial point process of rectangles. For each rectangle a data energy was defined, correlating possible rectangle hypothesis with the data. A regularizing energy acting on the spatial pattern of rectangles was incorporated, favoring for instance alignments between buildings. This model is robust with respect to the type and the quality of the data, but fails to process very noisy data.

In this paper, we extend our previous work [6] by examining the possibility of dealing with such noisy data. Our previous work showed robustness, due to the type of prior used acting

Ariana - Joint research group CNRS / INRIA / UNSA -INRIA, 2004 route des Lucioles, BP 93, Sophia Antipolis, France. Tel : +33 492 387 857, Fax : +33 492 387 643, Email : mathias.ortner.dump@gmail.com, xavier.descombes@inria.fr, josiane.zerubia@inria.fr

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE

2

on the spatial patterns of extracted features. We show in this paper that the point process approach allows the fusion of different information. We focus on the example of homogeneous regions and linear discontinuities, and propose a model of interacting point processes of rectangles and segments. We extend the work we presented in [7] by providing details on both models and algorithms as well as results on different real data.

### C. Point processes in image processing

Point process models in image processing can be seen as a natural extension of Markov Random Field approaches (MRFs). In the early eighties, MRFs have been introduced in the computer vision community through the works of Besag [8] and Geman and Geman [9]. In a Markov Random Field representation, an image is modeled as a realization of a collection of random values associated with each pixel in the image. Although these pixel-based approaches proved to be powerful for the analysis of dense urban areas at medium resolutions (e.g. classification of textures [10]), the advent of HR data has strengthened the need for approaches amenable to the consideration of the geometrical nature of urban scenes.

Point process models - that can be considered as a part of the wider "stochastic geometry" field, allow the modeling of images as random configurations of geometric shapes and provide a natural setup for the inclusion of a prior knowledge on the spatial pattern of features. Such models were first used in image processing by A. Baddeley and M.N.M. van Lieshout in [11]. Further work has been performed by H. Rue [12], [13] and Green [14] while more complex applications, like road or building extraction, have been studied in [6], [15]–[18]. Different ideas have been explored by Grenander, Miller and Sivrastava under the name of "pattern theory" (see [19] and reference therein), although the objects were not interacting.

### D. Outline

In Section II we provide a general discussion on point process models for automatic image feature extraction. We propose a generic model amenable to the inclusion of a prior knowledge on the pattern of features. In Section III, we then present the specific segment and rectangle point process models we adopted. Segments are used for detecting discontinuities, while rectangles are used for segmenting homogeneous areas. In Section IV, we provide an overview of the employed Reversible Jump Monte Carlo Markov Chain (RJMCMC) algorithm. We finally present and discuss results obtained on HR real data in Section V.

## II. POINT PROCESS MODELS FOR IMAGE FEATURE EXTRACTION

### A. Introduction to point processes

*1) Support, images and DEMs:* We model images as a *continuous* bounded set $K = [0, X_{1_{max}}] \times [0, X_{2_{max}}]$, and note $x = (c_1, c_2)$ a point of $K$. A Digital Elevation Model can therefore be described as a function associating a height with points of $K$. In the following, we note by $H$ such a function: $H : K \to [0, \infty[$.

*2) Random configuration of points:* A *configuration of points* $\mathbf{x}$ *in* $K$ (noted in bold) is a unordered set of points $\mathbf{x} = \{x_1, \ldots, x_{n(\mathbf{x})}\}$, where $x_i \in K$ and $n(\mathbf{x}) = \text{card}(\mathbf{x})$ denotes the number of points in the configuration. We note $\mathcal{C}$ the set of all possible finite configurations. Let consider a mapping from an abstract probability space $(\Omega, \mathcal{A}, \mathbf{P})$ to the set of configurations $\mathcal{C}$. Due to the finiteness of the considered configurations along with the boundedness of $K$, the $\sigma$-algebra associated with $\mathcal{C}$ is well defined (see [20] for details.) A *point process* $\mathbf{X}$ of points in $K$ is a measurable mapping that associates with an event $\omega \in \Omega$ a configuration of points $\mathbf{X}(\omega) = \{x_1, \ldots, x_n, \ldots\}$ $x_i \in K$. Accordingly, a point process is a random variable whose realizations are random configurations of points.

*3) Poisson point process:* The most random point process (in the entropy sense) is the *Poisson point process*. Let $\nu(.)$ be a positive measure on $K$. A Poisson point process $\mathbf{X}$ with intensity $\nu(.)$ verifies the following properties:

- for every Borel set $A \subset K$, the random variable $N_{\mathbf{X}}(A)$, giving the number of points of $\mathbf{X}$ falling in the set $A$, follows a discrete Poisson distribution with mean $\nu(A)$ (i.e. $\mathbf{P}(N_{\mathbf{X}}(A) = n) = e^{-\nu(A)} \frac{\nu(A)^n}{n!}$),
- and for every finite sequence of non intersecting Borelian sets $B_1, \ldots, B_p$ the corresponding random variables $N_{\mathbf{X}}(B_1), \ldots, N_{\mathbf{X}}(B_p)$ are independent.

Poisson point processes are usefull in our setup due to their analog role to Lebesgue measures on $\mathbb{R}^d$. As we detail it later, it is indeed possible to define point processes by their density with respect to the distribution of a reference Poisson point process.

*4) Marked point process:* The configurations of points described so far only include simple points of $\mathbb{R}^2$. To describe random configurations of geometrical objects, random marks are added to each point.

For instance, let consider the following mark set $M^r = ] -\frac{\pi}{2}, \frac{\pi}{2}] \times [L^r_{\min}, L^r_{\max}] \times [l^r_{\min}, l^r_{\max}]$. Noting by $x$ elements of $S^r = K \times M_r$, we consider the following parameterization describing rectangles $x = (c_1(x), c_2(x), \theta(x), L(x), l(x))$, where $(c_1(x), c_2(x))$ stands for the center position, $\theta(x)$, for the orientation, $L(x)$ for the length and $l(x)$ to the width of the rectangle $x$. A marked point process $\mathbf{X}$ of rectangles is a point process on $S^r = K \times M^r$ [1]. This parameterization is illustrated on Figure 1.

Similarly, we define a mark set $M^s$ to describe segments $M^s = [-\frac{\pi}{2}, \frac{\pi}{2}] \times [L^s_{min}, L^s_{max}]$. We note $y$ an element of $S^s = K \times M^s$ representing a segment $y \in S_s$, $y = (c_1(y), c_2(y), \theta(y), L(y))$, and use $\mathbf{Y}$ to denote a point process of segments.

Finally, we note $\mathcal{C}_r$ and $\mathcal{C}_s$ the sets of finite configurations of rectangles and segments.

*5) Density of a spatial point process:* An attractive feature of spatial point processes is the possibility of defining a point process distribution by its *probability density function* (pdf). A Poisson point process can indeed play the analog

---

[1]There is actually a further requirement that the restriction of $\mathbf{X}$ to $K$, noted $\mathbf{X}_{|K}$, should also be a point process on $K$. In our case, this technical condition on the measurability of the mapping $\mathbf{X}$ is satisfied since the sets $K$ and $S^r$ are bounded, see [20] for details.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE

3



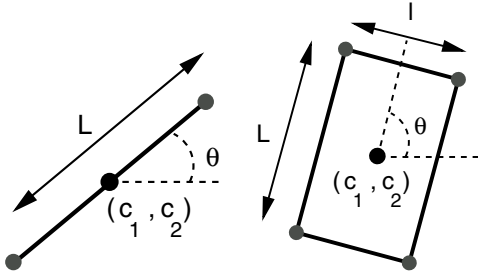Fig. 1.  Parameterizations of segments and rectangles

role to Lebesgue measure on $\mathbb{R}^d$. Consider the distribution $\mu(.)$ of a Poisson point process defined by its non atomic intensity measure $\nu(.)$ and a mapping $h(.)$ from the space of configurations of points $\mathcal{C}$ to $[0, \infty[$. We consider the function $Z(\mu, h)$ defined as $Z = \int_{\mathcal{C}} h(\mathbf{x}) d\mu(\mathbf{x})$. If $Z < \infty$, the function $h(\mathbf{x})/Z(\mu, h)$ can be seen as the density of a point process $\mathbf{X}$ with respect to the reference Poisson process (see [20]).

For instance, assume that $h(\mathbf{x}) = \prod_{i=1}^{n(\mathbf{x})} \beta(x_i)$ where $\beta(.)$ is an *intensity* function from $S$ to $]0, \infty[$. A point process $\mathbf{X}$ defined by this density turns to be a Poisson point process with intensity

$$\nu'(A) = \int_A \beta(u) d\nu(u). \tag{1}$$

In this simple case, the probability density function $h(.)/Z(\mu, h)$ allows a change of intensity measure. This example actually belongs to the more general class of *exponential families*. Let $t(.)$ be a mapping from $\mathcal{C}$ to $\mathbb{R}^k$. It is possible to describe a class of point process densities by using a parameter $\theta \in \mathbb{R}^k$ together with the scalar product $< ., . >$ and setting $h(\mathbf{x}) = e^{-<\theta, t(\mathbf{x})>}$. Of course, the density is well defined if and only if $Z(\mu, h) < \infty$. In this work we use such models to introduce a density where points are not independent like in the Poisson case but are correlated by means of interaction energies.

*6) Estimator and MCMC:* In [21] we presented an MCMC algorithm generating samples of a point process $\mathbf{X}$ defined by an unnormalized density $h(.)$ along with a reference Poisson point process distribution. The obtained algorithm produces a Markov Chain $(X_t)_{t \geq 0}$ ergodically converging to the distribution of $\mathbf{X}$.

The procedure permits the computation of Monte Carlo values. Another possibility is to use the sampler within a *simulated annealing* framework providing a global maximum of the density $h(.)$ as described in [22]. The estimator obtained is consequently the maximum density estimator $\hat{\mathbf{x}} = $ Argmax $h(.)$ The algorithm is detailed in Section IV.

*7) A short example : the Strauss process:* Consider a process of disks for wich $h(\mathbf{x}) \propto \exp(-\theta t(\mathbf{x}))$, where $t(\mathbf{x})$ counts the number of disk intersections within the configuration $\mathbf{x}$. For $\theta > 0$, overlapping disks are penalized. This model was used by Strauss in [23] to model patterns of trees. The sampling algorithm is based on an iterative process that tries to add or remove points in the configuration. The addition of a disk in the neighborhood of others will likely be rejected for $\theta$ large enough, while the deletion of an interaction will be favored. After enough iterations corresponding to a Markov

Chain convergence, the process will favor configurations of disks without overlapping pairs.

### B. Configurations of objects, image and energy

In this section, we define a suitable class of densities for the extraction of features in Digital Elevation Models. As mentioned previously, we focus on two types of elements: segments and rectangles. Our goal is to use the segments for detecting discontinuities and the rectangles to detect rectangular homogeneous areas standing for buildings. We therefore need to define a density making the objects fit the data (homogeneity for rectangles, discontinuity for segments) as well as favoring some patterns (connections of segments, interactions between segments and rectangles). In this section, we focus on generic modeling issues while the specific models are detailed in section III.

Let consider a point process $\mathbf{X}$ of features (e.g. either segments or rectangles). We recall here the strategy we adopted in [6] to define a suitable unnormalized density $h(.)$. In image processing, two main types of models are generally used. The first approach, *Bayesian modeling*, is known to be powerful but requires to exhibit a likelihood function describing the distribution of an image $I$ for a fixed configuration $\mathbf{x}$ of features (conditional distribution). The second approach uses the Gibbs distribution associated with a suitable energy function - as we describe it later. In our framework, the Bayesian approach would necessitate to accurately describe the distribution of heights in every pixel of the DEM for a fixed configuration of elements. This is a hard task, since by definition a feature only contains a small part of the image information. Examples of Bayesian point process models are given in [14] or [12]. They rely on the use of foreground and background models of grey level distribution. In [24] we proposed a Bayesian model for building extraction based on a point process of 3D buildings. The computational price for having a complete conditional description of the image is too heavy as it requires too many random parameters.

In this work, we use the second class of widely used models and define a density under its Gibbs form

$$h(.) = \frac{1}{Z} e^{-U(\mathbf{x})} \quad U(\mathbf{x}) = U_{int}(\mathbf{x}) + \rho U_{ext}(\mathbf{x}). \tag{2}$$

The energy is divided into two parts. The internal field $U_{int}(\mathbf{x})$ favors specific spatial structures in the configuration $\mathbf{x}$, while the external field $U_{ext}(\mathbf{x})$ quantifies the quality of the configuration with respect to the data. The positive parameter $\rho$ allows the tuning of the relative weights between the two terms.

The simplest way of specifying a data term is to expand it as a sum over the objects in a configuration

$$U_{ext}(\mathbf{x}) = \sum_{u \in \mathbf{x}} U_d(u). \tag{3}$$

The definition of a data term for the whole configuration is thereby reduced to the definition of a data term for one object which is faster from a computational point of view. The mapping $U_d(.)$ from $S = S^r \cup S^s$ to $\mathbb{R}$ quantifies the relevance of an object with respect to the data. Note that a

4

Bayesian modeling with respect to the data image $I$ would result in a log-likelihood $-U_{ext}(\mathbf{x}|I)$ that has no reasons to be expandable over the objects and that would be heavy from a computational point of view.

If the data energy of an object $u$ is negative ($U_d(u) \leq 0$), we say that the object is *attractive*. Care is needed to avoid superpositions of points. From equation (3) it is obvious that if $U_d(u) \leq 0$, then successive additions of clones of $u$ decrease the overall data energy $U_{ext}(\mathbf{x} \cup u \cup u) \leq U_{ext}(\mathbf{x} \cup u) \leq U_{ext}(\mathbf{x})$. As a consequence, a repulsive term avoiding such superpositions is needed, and we eventually add an *exclusion* term $U_{excl}(\mathbf{x})$ to the energy

$$U(\mathbf{x}) = \rho \sum_{u \in \mathbf{x}} U_d(u) + U_{int}(\mathbf{x}) + U_{excl}(\mathbf{x}), \qquad (4)$$

such that $U(\mathbf{x} \cup u \cup u) > U(\mathbf{x} \cup u)$, for all $(u, \mathbf{x}) \in S \times \mathcal{C}$.

### C. Internal field

The purpose of the internal field is to favor pre-defined patterns of objects. In the specific case of segments and rectangles, our purpose is to favor paving patterns of rectangles, connections of segments, and completion between rectangles and segments.

The simplest Gibbs point process with non-independent points is the *Strauss process*. Let consider a point process $\mathbf{X}$ in $K$ describing random configurations of simple points $\mathbf{X}(\omega) = \{x_1, \ldots, x_{n(\mathbf{X}(\omega))}\}$ with $x_i = (c_1^i, c_2^i) \in \mathbb{R}^2$.

The Strauss process is based on the energy $U(\mathbf{x}) = \gamma s(\mathbf{x})$. The function $s(\mathbf{x})$ counts the number of pairs of points that are closer than a parameter $\delta$, while $\gamma > 0$ is a real parameter tuning the importance of the interaction term. In view of Equation (2) the Strauss process penalizes configurations with too many close points. Note that the process can alternatively be seen as a process of disks with radius $\delta/2$, and $s(\mathbf{x})$ counting the number of intersections between disks.

Strauss processes were originally introduced [23] for modeling patterns of trees, and the repulsive nature of the interaction was obviously suitable to that particular application. In our case, we would like to favor some clusters (e.g. segment connections). An naive solution would be to take a similar model and set $\gamma < 0$. However, in that case, the process is not defined ($Z = \infty$). The behavior of $s(\mathbf{x})$ is indeed in $n(\mathbf{x})^2$ (see [25] or [21] for more details).

*1) Definitions:* For a given relation $\sim$ on $S$, let note $\mathcal{R}(\mathbf{x})$ the set of interacting couples of $\mathbf{x}$

$$\mathcal{R}(\mathbf{x}) = \{(u, v) \ : \ u \in \mathbf{x}, \ v \in \mathbf{x}, \ u \neq v \ u \sim v\}. \quad (5)$$

The relation $\sim$ can be symmetric, but it is not required (see [26] for theoretical ramifications regarding non-symmetric neighbor relations). We define the neighborhood $\mathcal{N}(\mathbf{x}, u)$ of a point $u$ in $\mathbf{x}$ as the set of points in $\mathbf{x}$ that are in relation with $u$

$$\mathcal{N}(u, \mathbf{x}) = \{v \in x \ : \ u \sim v\}. \qquad (6)$$

Let consider the function $V(\mathbf{x}, u) = \mathbf{1}(\mathcal{N}(u, \mathbf{x}) \neq \emptyset)$ which is null only if $u$ has no neighbors in $\mathbf{x}$. This function is included in the model in order to favor or penalize the presence of an interacting pair of points.

We are also interested in ordering interactions with respect to a quality function (see [18]). We suppose that for each type of interaction, a function $\Psi(.,.)$ from $S \times S$ to $[-1, 1]$ quantifies the quality of the interaction between interacting objects (intuitiveley $\Psi(u, v)$ close to 1 (resp. $-1$) means that the interaction is good (resp. bad) while $\Psi(u, v) = 0$ whenever $(u, v)$ do not interact).

Care is needed with the incorporation of those $\Psi$ functions in the configuration energy. Similarly to the attractive Strauss model, summing the $\Psi$ values over the interacting couples would result in a non integrable $h(.)$. Again, since an object is possibly involved into several interactions, the number of interacting couples can evolve with a $n(\mathbf{x})^2$ behavior. As a consequence, we propose to compute for each object the maximum reward value

$$W(\mathbf{x}, u) = \begin{cases} \max_{v \in \mathcal{N}(\mathbf{x}, v)} \Psi(u, v) & \text{if } \mathcal{N}(\mathbf{x}, u) \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

The function $W(u, \mathbf{x})$ is the reward function of the best interaction among those involving $u$. Note that for a repulsive interaction it might be better to compute the worst one (minimizing $\Psi$).

*2) Local energies:* We define the *local energy* of an object $u \in \mathbf{x}$ associated with a specific interaction ($i$-th) to be a linear combination of the corresponding functions $V$ and $W$,

$$u \in \mathbf{x}, \qquad U_{loc}^i(\mathbf{x}, u) = -\left(a^i V^i(\mathbf{x}, u) + b^i W^i(\mathbf{x}, u)\right). \tag{7}$$

Here $a$ and $b$ are two real parameters tuned to favor ($a, b > 0$) or penalize ($a, b < 0$) patterns of interest; $a$ favoring the presence of an interacting object and $b$ weighting the quality of the best interaction involving $u$.

We finally define the total interaction energy of the configuration as the sum of local energies over the objects

$$U^i(\mathbf{x}) = \sum_{u \in \mathbf{x}} U_{loc}^i(\mathbf{x}, u). \tag{8}$$

*3) Generalization:* When using several interactions $\sim^1, \ldots, \sim^k$, we naturally extend the model by summing Equation (8) over the different interactions

$$U_{int}(\mathbf{x}) = \sum_{i=1}^{k} U^i(\mathbf{x}) = \sum_{u \in \mathbf{x}} \sum_{i=1}^{k} U_{loc}^i(\mathbf{x}, u) \tag{9}$$

$$= -\sum_{u \in \mathbf{x}} \sum_{i=1}^{k} [a^i V^i(\mathbf{x}, u) + b^i W^i(\mathbf{x}, u)]. \tag{10}$$

The internal field accordingly evolves linearly with the number of points in the configuration. Combinatorial problems therefore disappear and the process is well defined. Another advantage is that the balance between the external field (data term) and the internal field is eased. An important benefit of this interaction model is its scale invariance. The weights $a^i, b^i$ actually do not depend on the size of the considered area, provided that the density of objects to be detected is constant, a point we discuss in Section IV.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE

5

## D. Exclusion interaction

Although the attractive interaction model described previously solves the issue of the model integrability, an exclusion term penalizing redundant objects is needed. We already provided an immediate justification for this need through the discussion on the data term and the definition of attractive objects. It should also be noticed that attractive interactions enhance this need. Without an exclusion term, the density maximum can consist of an infinite accumulation of points, even if the density is integrable. A last reason for adding a repulsive interaction comes from a condition to get the algorithm convergence (see [21]) which requires the energy variation induced by adding a point to be uniformly bounded.

These three reasons underline the need for an exclusion term. Since we consider geometrical objects, the simplest exclusion interaction we can use is the *intersection relation* $u \sim^{excl} v$ if and only if $\text{Surf}(u) \cap \text{Surf}(v) \neq \emptyset$ where we note by $\text{Surf}(u)$ the silhouette of the geometrical object $u$.

To incorporate this interaction as a repulsive one, we use a simple model, homogeneous to the general interaction model of Equation (7)

$$U_{excl}(\mathbf{x}) = -a^{excl} \sum_{u \in \mathbf{x}} V_{excl}(\mathbf{x}, u) \qquad a^{excl} < 0. \quad (11)$$

Here $|a^{excl}|$ is taken large enough so that it is impossible to have redundant objects in the maximizing configuration. However, due to the linear property of our attractive model, the exclusion weight does not need to be too large, resulting in good mixing properties of the algorithm.

## E. Data term

We provide here a generic analysis of the data term $U_d(u)$ associated with an object $u$.

*a) Attractive objects:* A data term $U_d : S \rightarrow \mathbb{R}$ partitions the set $S$ into the set of attractive objects $U_d \leq 0$ and its complement (repulsive objects). Since the density $h$ is maximized by a simulated annealing procedure, a repulsive object cannot be part of the resulting configuration, except if attractive interactions force its presence. For the sake of clarity, we avoid that phenomenon by considering attractive interactions among attractive objects only. We note $\gamma_1^r \subset S^r$ the set of attractive rectangles and $\gamma_0^r$ its complement. Similarly the set of attractive segments is noted $\gamma_1^s$ while the set of repulsive segments is noted $\gamma_0^s$.

*b) Attractive object ordering:* Another interesting feature of the function $U_d$ is that it allows ordering the objects. We separate two cases. Among the repulsive objects, we use this function to favor repulsive objects "close to" the set of attraction $\gamma_1$ whereas for attractive objects we use $U_d(.)$ to favor the specific locations that fit the data the best.

## III. SPECIFIC MODEL DEDICATED TO THE ANALYSIS OF DEMS

The purpose of this section is to present a specific model for the analysis of Digital Elevation Models. We define the rectangle and segment processes in terms of their internal field and data terms. We also present an interaction term to make both processes interact.

## A. Segments and discontinuities

We enumerate here the different parts composing the segment model. We first explicit the data term, whose goal is to make the linear network match meaningful discontinuities of the DEM. We then describe the attractive connection interactions used along with the corresponding repulsive term. We conclude this section by presenting a simulated example of the action of the internal field on the segment point process.

*1) Configurations of segments:* As mentioned in the previous section we consider a segment space $S^s$, product of the image $K$ and the segment mark space $S^s = K \times M^s$, where $M^s =] -\frac{\pi}{2}, \frac{\pi}{2}] \times [L_{min}, L_{max}]$. Points in $S^s = K \times M^s$ are segments parameterized by their center in the image together with their orientation and length. A marked point process $\mathbf{Y}$ on $S^s$ describes random configurations of segments $\mathbf{Y}(\omega) = \{y_1, \ldots, y_{n(\mathbf{Y}(\omega))}\}$ with $y_i \in S^s$, for all $i \in \{1, \ldots, n(\mathbf{y})\}$.

*2) Data term:* Our purpose is to use linear networks for detecting significant discontinuities on the DEM. Segment networks have previously been used in the context of road detection ( [16], [15]). The aim of the data term is to quantify the relevance of a segment hypothesis with respect to the DEM. Since we want to process various type of DEMs, we propose to use the generic discontinuity detector that we presented in [6].

*a) Discontinuity filter:* Figure 2 provides a graphical explanation of our approach. For a given segment hypothesis, we consider slices taken orthogonally to the segment direction on the DEM. For each of these slices, the discontinuity filter detects discontinuity locations.

We use a dedicated filter. This filter is based on the accumulation of large enough gradient values, and permits the analysis of smooth data (e.g. Laser measurements), as well as sharper DEMs (e.g. provided by optical stereovision). This detector relies on the following steps. First, on each profile, the points such that the local gradient is larger than a parameter $\nabla H_{\min}$ are pre-selected. Second, the pre-selected gradients are accumulated, and selected if the accumulation is greater that $\Delta H_{min}$. Finally, the discontinuity closest to the segment is kept as the important point. A morphological opening step is also performed, using a linear element of size $l_{sel}$. A complete description of the filter is given in [27]. An important point is that this filter is implicitly directional as the slice analysis is performed orthogonally to a segment hypothesis.

*b) Reward functions:* We detail here how the selected discontinuities are incorporated into the data term. As already discussed in the previous section, the data term $U_d(.)$ enables the discrimination between attractive and repulsive objects.

For a given segment hypothesis $u$ on which the discontinuity analysis has been performed, we compute the length of the detected discontinuity $Lg(u)$ as illustrated by the Figure 2(c). This length is given by the ratio of selected discontinuities, that are at a distance to the segment smaller than a parameter $\delta r$, multiplied by the length of the segment. We also compute a moment value, $\hat{m}(u)$, as illustrated by Figure 2(d). This moment is the average squared distance between the detected discontinuities and the segment.

*c) Final data term:* We define the set of attractive segments as the set of segments such that the ratio of detected

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE

6



(a) A segment hypothesis (in black) on a Digital Elevation Model and the considered slices, orthogonal to the segment.

(b) We use a dedicated filter to detect on each slice the closest meaningfull discontinuity to the segment.

(c) The detected discontinuity length corresponds to the length of the gray rectangle.

(d) The discontinuity moment is the average squared distance between the selected discontinuities and the segment.
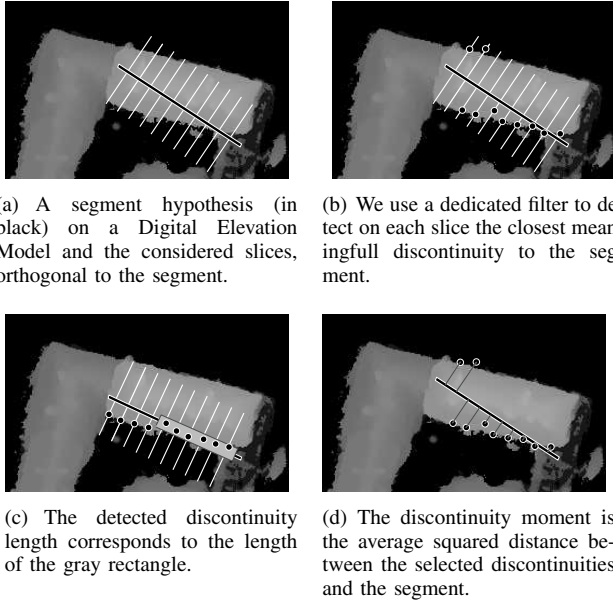
Fig. 2. Illustration of the steps used to compute the data energy of a segment hypothesis. First, orthogonal slices are extracted from the DEM (see 2(a)). Then, using a dedicated filter, on each profile the closest meaningfull discontinuity is detected (see 2(b)). The data energy is based on the length of detected discontinuity, which corresponds to the ratio of discontinuities that are close enough to the segment multiplied by its length (see 2(c)); as well as the average squared distance between the discontinuities and the segment, a value that we call discontinuity moment (see 2(d)).

discontinuities is greater than a fixed threshold $\eta \in [0, 1]]$

$$\gamma_1^s = \{u \in S^s \qquad \text{s.t.} \quad Lg(u) \geq \eta L(u)\}$$

In practice, we take $\eta = 90\%$. As a consequence, the set of attractive segments is the set of segments lying on discontinuities. We also define a reward function $j_{seg} : S^s \rightarrow \mathbb{R}$ associated with a segment

$$j_{seg}(u) = \frac{1}{2} \frac{Lg(u)}{L_{max}} + \frac{1}{2}(1 - \hat{m}(u)).$$

This reward function favors segments with a large discontinuity length and fitting the discontinuities well. Therefore we get

$$U_d^s(u) = \left(-j_{seg}(u) * \mathbf{1}_{\gamma_1^s}(u) + 0.1 * (2 - j_{seg}(u))\mathbf{1}_{\gamma_0^s(u)}\right)$$

This data term favors the attractive segments and leads elements of $\gamma_0^s$ towards the set of attractive elements $\gamma_1^s$ as it associates low values with segments in good correspondence with the discontinuities of the DEM

*3) Internal field for the segment:* We describe in this section the internal field designed to favor configurations of connected segments. We consider two different connection cases depending on the angle between the segments. We promote cases where the connection happens between aligned or orthogonal segments.

*a) Connections:* Connection interactions for segment processes has been considered in the works of [16] and [28]. We use similar geometrical interactions, but our generic model of attractive interactions presented in section II-C allows an easier incorporation of a quality term.

We introduce a connection relation $\sim_{conn}$ depicted by Figure 3. We identify by $E_1(u)$ and $E_2(u)$ the extremities of the segment $u$. We consider that two segments are connected if the distance between their extremities is less than a parameter $\delta E_{max}$ : $u \sim_{conn} v$ iff $\min\{\|E_i(u) - E_j(v)\| \quad i, j \in \{1, 2\}^2\} \leq \delta E_{max}$

*b) Restrictions:* We actually consider two different types of connections depending on the angle between the connected segments, as illustrated in Figure 3. We therefore append the following two conditions on the difference of angles: $|\theta(u) - \theta(v)| \leq \delta\theta_{max}$ (modulo $\pi$) and $|\theta(u) - \theta(v) + \frac{\pi}{2}| \leq \delta\theta_{max}$ (modulo $\pi$).



(a) A connection between aligned segments.

(b) A connection between orthogonal segments.

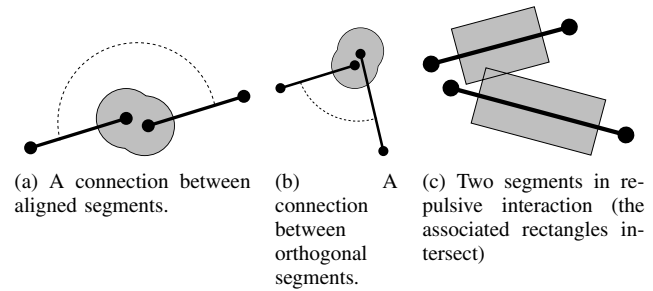(c) Two segments in repulsive interaction (the associated rectangles intersect)

Fig. 3. Interactions between segments. Two segments are connected if the distance between their extremities is small enough.. Two types of connections are actually distinguished: connections with a flat angle (see 3(a)) and connections with a right angle (see 3(b)). A repulsive intersection relation (see 3(c)) is added to avoid overlapping segments.

We finally add a further restriction, and consider only connections between attractive segments. Since a segment has two extremities, we consider two relations of each type per segment. The different resulting mathematical expressions are detailed in [29].

*4) Energy Model:* In order to favor good connections we introduce the following real valued function

$$\varpi : \quad \mathbb{R}^2 \quad \rightarrow \quad [0, 1]$$
$$(x, x_{max}) \quad \rightarrow \quad \frac{1}{x_{max}^2}\left(\frac{1 + x_{max}^2}{1 + x^2} - 1\right) \quad , |x| \leq x_{max}. \tag{12}$$

This function verifies $\varpi(0, x_{max}) = 1$ and $\varpi(x_{max}, x_{max}) = 0$. In the case of the flat connection relation $\sim^{conn.al._1}$ we adopt the following reward function

$$\Psi(u, v) = \frac{1}{2}\varpi(\delta E(u, v), \delta E_{max}) + \frac{1}{2}\varpi(\delta\theta, \delta\theta_{max}).$$

where $\delta E$ and $\delta\theta$ respectively represent the distance between the suitable extremities and the suitable angle difference. We therefore favor precise connections (the smaller the distance between the extremities, the better) as well as flat connections. The quality function associated with the orthogonal connection is similar and favors angle differences close to $\pi/2$.

We use the generic model presented in section II-C. This model relies on four real parameters, $a^{\text{conn. al.}}, b^{\text{conn. al.}}, a^{\text{conn. orth.}}, b^{\text{conn. orth.}}$. As we want the connections to be favored, these parameters are taken positive. Note that since we have 4 connection relations, we should consider 4 couples $(a, b)$. However, since the connections by one or another extremity is symmetric (relation $\sim^{\text{conn. al.}_1}$
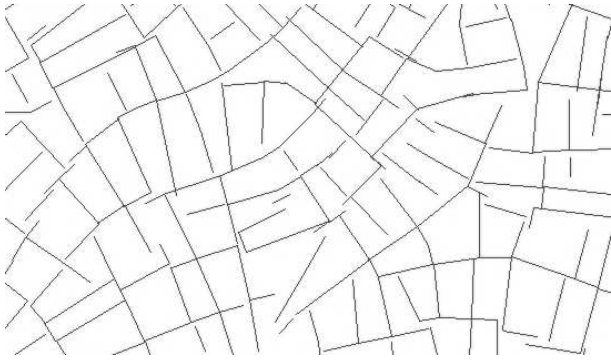
Fig. 4. Simulation result of the internal field acting on the segment process. This simulation has been obtained using $\gamma_1^s = S^s$.



(a) A mask of points, associated with each rectangle, is made of a set of central points and four bands along the sides.

(b) The four bands are used to obtain a ground height estimate taken as the lowest average height among the four bands.

(c) The extruded ratio is given by the ratio of inside points higher than the ground height estimate augmented by a minimum height (points in white).

Fig. 5. To compute the data energy of a rectangle we use a mask of points associated with the rectangle (see 5(a)). The four lateral bands gives four mean DEM height (see 5(b)). The lowest mean gives a local estimate of the ground height. We then compute a volume ratio (see 5(c)) that corresponds to the ratio of inside points that are higher than the ground height augmented by a minimum height, set by the user. We also compute the standard deviation of these points.

equivalent to $\sim^{\text{conn. al.}2}$) we set $a^{\text{conn. al.}1} = a^{\text{conn. al.}2} = a^{\text{conn. al.}}$ and proceed similarly for the other parameters.

*5) Exclusion term.:* Following the discussion in section II-D on the necessity of introducing a repulsive interaction to avoid accumulations of attractive points, we incorporate within the model an intersection interaction. For each segment, we define an embedding rectangle (see Figure 3(c)) and set the exclusion interaction as the intersection between the corresponding rectangles.

*6) Result example:* We present on Figure 4 a simulation result of the prior model. The internal field favors connected networks of lines and both flat and orthogonal connections.

### B. Rectangles and homogeneity

In this section we detail the rectangle process. After recalling the object space, we describe the data term as well as the internal field. We then show simulated results of the model.
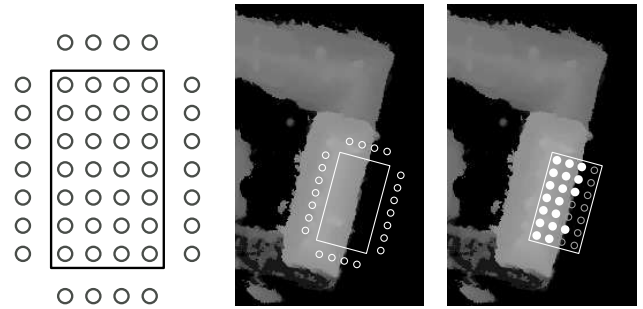
*1) Configurations of rectangles:* We define a process of rectangles $\mathbf{X}$ whose points represent rectangles. The object space is $S^r = K \times M^r$. $M^r = ] - \frac{\pi}{2}, \frac{\pi}{2}] \times [L_{min}^r, L_{max}^r] \times [l_{min}^r, l_{max}^r]$.

*2) Rectangles and DEM:* The goal of the data term is to make the rectangles fit extruded areas of the Digital Elevation Model. For that purpose we use a mask of points presented on Figure 5(a). This mask is made of a set of points inside a rectangle along with four bands around it.

*a) Local ground height:* We compute a local ground height estimate $\hat{H}_g(u)$ from the four bands. We take the minimum of the four means of heights. This procedure provides a local estimate of the ground height. This point is important, as dense urban areas often exhibit different "ground levels".

*b) Extruded ratio:* We define the volume ratio $\bar{v}(u)$ as the percentage of points inside the rectangle that are higher than the ground height estimate augmented by a minimum height $(\hat{H}_g(u) + H_{min})$.

*c) Surface ratio:* Another function we use is the ratio between the area of the rectangle and the maximum possible area. $\bar{s}(u) = l(u) * L(u)/(L_{max}^r * l_{max}^r)$. This function is included in the data term to favor large rectangles. It should be noted that a term favoring large rectangles could be considered as a part of the internal field.

*d) Standard deviation ratio:* We compute the standard deviation $\sigma(u)$ of grey levels among the points of the mask that are higher than the minimum height $\hat{H}_g(u) + H_{min}$. The standard deviation ratio is given by $\bar{\sigma}(u) = \max\{0, 1 - \sigma(u)/\sigma_{\max}\}$. This term obviously favors an homogeneous distribution of gray levels inside the rectangles. Note that this term relies on an additional parameter $\sigma_{\max}$, which in practice we take as $\sigma_{\max} = 10m$.

*3) Data term:*

*a) Attractive objects:* We define the set of attractive rectangles as the set of rectangles for which the extruded ratio is large enough, i.e. $u \in \gamma_1^r$ iff $\bar{v}(u) \geq v_{min}$. In practice, we use $v_{min} = 90\%$ and $H_{min} = 4m$.

*b) Data function $U_d$:* We consider the following reward function associated with a rectangle $j_{rect}(u) = \bar{v}(u) * \bar{s}(u)/2 + \bar{\sigma}(u)/2$. By multiplying the extruded and surface ratios, we actually obtain the surface of the rectangle that is above $\hat{H}_g(u) + H_{min}$ up to proportionality. This function favors large homogeneous rectangles. We finally end up with the following data term

$$U_d^r(u) = \left( -j_{rect}(u) * \mathbf{1}_{\gamma_1^r}(u) + 0.1 * (2 - j_{rect}(u))\mathbf{1}_{\gamma_0^r}(u) \right)$$

Again, this function $U_d^r(.)$ discriminates between attractive and repulsive objects and favors the best objects, in the sense of the homogeneity of gray levels.

*4) Internal field for the rectangle process:* In a town, buildings are usually aligned. Hence, we design an interaction that favors such alignments. Figure 6(a) presents an example of alignment. This model is similar to the model we propose in [6].

*a) Alignment interaction:* Denoting $\delta C(u, v)$ the distance between appropriate corners (see details in [6]) and $\delta\theta(u, v)$ the angle difference between the two rectangles (modulo $\pi$), we define the first interaction by the three following

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

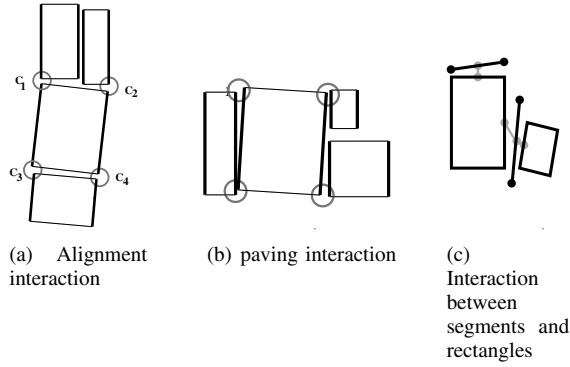IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE

8

Fig. 6.   Some attractive interactions. a) example of an alignment interaction, b) Illustration of a paving interaction (alignment along the other axis) and c) illustration of an interaction between rectangles and segments.

conditions

$$u \sim^{al_1} v \iff \begin{cases} \delta C(u,v) \leq \delta C_{\max} \\ \delta\theta(u,v) \leq \delta\theta_{\max} \\ (u,v) \in \gamma_1^2. \end{cases} \quad (13)$$

Because a rectangle has four corners, we actually define four different alignment interactions that we note $\sim^{al_1}, \ldots, \sim^{al_4}$. We only detect alignments between two attractive objects (belonging to $\gamma_1^r$). A rectangle can be in relation with a neighbor by two corners in which case we consider that they are related by two different interactions. The associated reward function evaluated on a couple of related points is:

$$\Psi(u,v) = \frac{1}{2}\varpi\left(\delta\theta(u,v), \delta\theta_{\max}\right) + \frac{1}{2}\varpi\left(\delta C(u,v), \delta C_{\max}\right). \quad (14)$$

This reward function is important: the goal is not only to promote the presence of alignments, but also to favor alignments of good quality. We finally define the internal field

$$U_{int}(\mathbf{x}) = \sum_{u \in x} a^{al} \sum_{i=1}^{4} V^{al_i}(\mathbf{x}, u) + b^{al} \sum_{i=1}^{4} W^{al_i}(\mathbf{x}, u). \quad (15)$$

Note that although 4 alignment relations are considered, we only use two parameters $a$ and $b$ to limit the number of parameters.

*b) Paving interaction:* We also introduce a second type of relation which favors parallel rectangles that are located side by side as illustrated by Figure 6(b). This interaction is essentially introduced in order to favor clean arrangements of buildings. This relation is defined similarly to the alignment relation previously described. The reward function $\Psi$ is also defined in a similar way. It leads to interactions $\sim^{pav_1 1}$ to $\sim^{pav_4}$.

*5) Results:* An illustrative example of a simulation of the internal field is presented on Figure 7.

### C. Coopeeration between segments and rectangles

We present in this section an interaction between segments and rectangles whose purpose is to merge the two processes in order to obtain a joint analysis of the data based on both discontinuities and extruded area detection.
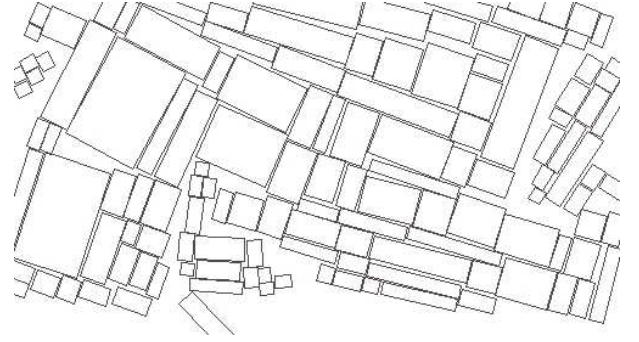


Fig. 7.   Simulation result showing the internal field of the rectangle process.

*1) Union of point processes:* We formally denote configurations of rectangles and segments as realizations of a process $\mathbf{Z}$, union of the processes $\mathbf{X}$ of rectangles and $\mathbf{Y}$ of segments $\mathbf{Z} = \mathbf{X} \cup \mathbf{Y}$.

Alternatively, the point process $\mathbf{Z}$ can be seen as a marked point process on $S^{sr} = K \times ((\{s\} \times M^s) \cup (\{r\} \times M^r))$. An object $z$ is accordingly described by a point, an indicator $o \in \{s, r\}$, and the relevant marks. This alternative representation is useful from the algorithmic point of view, as it allows us to extend the convergence results we obtained in [6] and [21] to the case of a process of rectangles and segments.

*2) Interaction between segments and rectangles:* We define an interaction term between rectangles and segments in order to favor the coherence between the two networks of objects. We use a test on the angle difference, as well as a test on the distance between the segment center and the closest side of the rectangle. The associated limit parameters are noted $\delta d_{\max}$ and $\delta\theta^{rs}$.

Similarly to the previously defined interactions, we consider only interactions among attractive objects. Note that a segment can be in relation with two rectangles (two sides) while a rectangle can be related to four segments, as illustrated by the Figure 6(c).

We end up with the following interaction energy

$$U_{coop}(\mathbf{z}) = \sum_{u \in z} a^{coop} V(\mathbf{z}, u) + b^{coop} W^{coop}(\mathbf{z}, u). \quad (16)$$

where $a^{coop}$, $b^{coop}$ are the interaction parameters, and $V, W$ the interaction functions defined similarly to the previous interactions.

*3) Resulting model:* Accounting for both the segment and rectangle processes result in the overall model

$$\begin{aligned} U^{rs}(\mathbf{z}) &= U^r(\mathbf{x}) + U^s(\mathbf{y}) + U^{inter}(\mathbf{z}) \\ &= \rho^r \sum_{u \in \mathbf{x}} U_d^r(u) + \rho^s \sum_{v \in y} U_d^s(v) + U_{int}^r(\mathbf{x}) + U_{int}^s(\mathbf{x}) + \ldots \\ &\quad + \ldots U_{excl}^r(\mathbf{x}) + U_{excl}^s(\mathbf{y}) + U_{coop}(\mathbf{z}) \end{aligned}$$

parameterized by $\rho^r$, $\rho^s$ together with the $a^i$ and $b^i$ covering the two types of segment connections, the two types of rectangle alignment relations, the two repulsive interactions and the cooperation interaction.

### IV. ALGORITHM

We present briefly in this section the algorithm used for optimizing the model described by Equation (4). Rather than

detailing technical issues that can be found in [25] or [21] we try to give a general flavor of the MCMC sampler we use.

### A. Sampling unnormalized densities

As stated in section II, we specify point process models by means of an energy $U(.)$ and a reference Poisson point process playing the analog role of real p.d.f. and Lebesgue measure for real random variables. The resulting distribution is denoted $\pi(.)$ and is known *up to a normalizing constant*. Due to the Markov Random Field approaches, it is now widely known in the image processing community that Gibbs or Metropolis-Hastings procedures allow the computation of Monte Carlo estimators, even if the normalizing constant of the target distribution is not known. Such algorithms are Monte Carlo Markov Chain samplers generating Markov chains that exhibit the desired properties (i.e. Harris recurrence, aperiodicity, $\pi$ invariance and ergodicity). As a consequence of these properties, the states of the Markov chain after a large enough time are distributed according to the desired distribution $\pi(.)$ and can be used to compute Monte Carlo values even if the samples are not independent, regardless of the starting point. Of course, mild conditions on the function to evaluate need to be fulfilled (e.g. it has to verify Lyapunov stability conditions, see [25]).

The point process case can be tackled similarly, by employing a Metropolis-Hastings like procedure. However, as it was pointed out by Green [30], when exploring state spaces of different dimensions, a correcting term has to be appended to the acceptance ratio when the Markov chain jumps between spaces of different dimensions.

As we now detail it, the Metropolis-Hastings-Green procedure applied to point processes is very similar to the usual Metropolis-Hastings update scheme. The transition kernel of the Markov chain is a two stage procedure. A possible new state is first randomly proposed by a perturbation kernel (*proposition step*) and then randomly accepted according to a suitable Bernoulli scheme (*acceptance step*). In the case of point processes, the current state is a configuration of geometrical objects. The random proposition can be for instance the translation or the rotation of one or several objects in the configuration or the addition/deletion of an item to/from the current configuration. In that case, the Green correcting factor is needed, since the dimension of the configuration is changed. The second step (random acceptance of the proposition) is the step that ensures the convergence of the Markov Chain to the desired distribution and the acceptance probability needs to be computed carefully.

The Metropolis-Hastings-Green algorithm applied to the specific case of point processes is known under the name of Geyer and Møller [31] algorithm by the point process community since they provided a proof of its convergence in 1994.

### B. Generic Structure

Suppose we consider a point process $\mathbf{Z}$ defined by its energy $U(.)$. Through the Gibbs relation, this energy leads to a density $h$ known up to a normalizing constant. This density

together with the distribution $\mu(.)$ of the reference Poisson point process defines the distribution $\pi(.)$ of $\mathbf{Z}$.

The Markov chain $(X_t)_{t \geq 0}$ is defined by a starting point $X_0 = \{\emptyset\}$ and a Markovian transition kernel $P(\mathbf{x}, .)$ corresponding to the conditional distribution of $X_{t+1}|X_t = \mathbf{x}$. It results in a Markov chain $(X_t)_{t \geq 0}$ on the space of finite configurations of points $\mathcal{C}$.

Of course, $P(., .)$ is designed in order to make the Markov Chain converge towards the desired distribution $\|P^n(\{\emptyset\}, .) - \pi(.)\|_{TV} \to 0$ where $\|.\|_{TV}$ notes the Total Variation norm (TV).

The Markov chain generated by the following algorithm satisfies this property. We actually have more accurate results, since we know that we can start from any configuration (Harris recurrence) and that the total variation tends to zero geometrically (geometric ergodicity), as detailed in [21].

*1) Algorithm:* The algorithm is based on a mixture of perturbation kernels $Q(., .) = \sum_m p_m Q_m(., .)$ where $\sum p_m = 1$ and $\int Q_m(\mathbf{z}, \mathbf{z}')\mu(d\mathbf{z}') = 1$. The algorithm iterates the following steps. Let the current state $X_t$ be $X_t = \mathbf{z} = \{z_1, \ldots, z_n\}$.

[1] Choose one of the proposition kernels $Q_m(., .)$ with probability $p_m(\mathbf{z})$ and

[2] sample $\mathbf{z}'$ according to the chosen kernel $\mathbf{z}' \sim Q_m(\mathbf{z}, .)$.

[3] Compute the Green ratio $R_m(\mathbf{z}, \mathbf{z}')$, function of the selected kernel $Q_m$, the original state $\mathbf{z}$ and the proposed new state $\mathbf{z}'$. The ratio $R_m$ is derived to make the Markov chain converge towards the desired distribution.

[4] The proposition is accepted ($X_{t+1} = \mathbf{z}'$) with a probability $\alpha_m(\mathbf{z}, \mathbf{z}') = min(R_m(\mathbf{z}, \mathbf{z}'), 1)$ and rejected otherwise ($X_{t+1} = \mathbf{z}$).

*2) Perturbation kernels:* The efficiency of the algorithm highly depends on the variety of possible transformations $Q_m(\mathbf{z}, .)$.

*a) Birth or death:* This kind of perturbation first chooses with probability $p_b$ and $p_d = 1 - p_b$ whether a point should be removed (death) or added (birth) to the configuration. If death is chosen, the kernel selects randomly one point $u$ in $\mathbf{z}$ and proposes $\mathbf{z}' = \mathbf{z} \setminus u$, while if birth is chosen, it generates a new point $u$ according to the uniform measure $|.|/|S^{sr}|$ and proposes $\mathbf{z}' = \mathbf{z} \cup u$. The birth or death kernel is necessary and sufficient to insure the convergence of the Markov chain towards the target distribution.

*b) Non jumping transformations:* Non jumping transformations are transformations that first select randomly a point $u$ in the current configuration and then propose replacing this point by a perturbed version $v$, $\mathbf{z}' = \mathbf{z} \setminus u \cup v$. Translation, rotation or dilation are examples of non jumping perturbations.

*c) Birth or death in a neighborhood:* We introduced this kind of transformation in [21]. The idea is to propose the removal or addition of interacting pairs of points with respect to one of the attractive relations such as the connection in the case of segments or alignment in the case of rectangles. This type of transformations increase the performance of the sampler since it follows the prior model (see [21].

*d) Green ratio:* With each of these proposition kernels a mapping $R_m(., .)$ from $\mathcal{C} \times \mathcal{C}$ to $(0, \infty)$ is associated. This value, named *Green ratio*, depends on the target distribution $\pi$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE

10



(a) Rotation of a rectangle.
(b) Rectangle dilation in one direction.
(c) Rectangle dilation in the second direction.
(d) Rectangle translation.

(e) Segment dilation.
(f) Segment rotation.
(g) Segment translation.
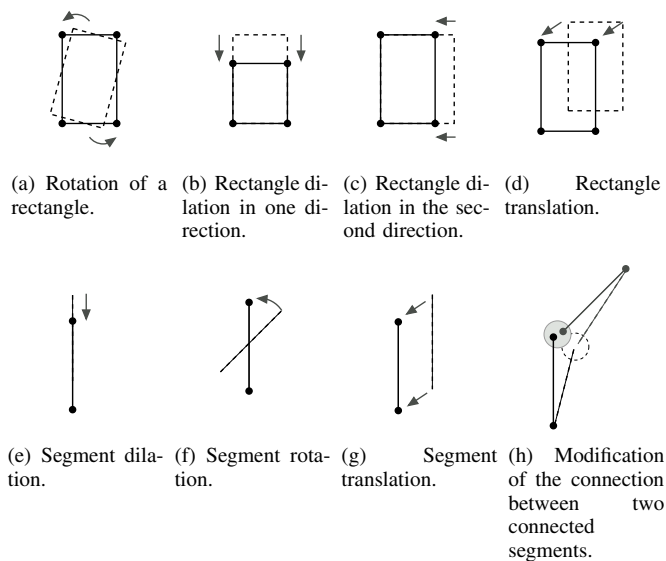(h) Modification of the connection between two connected segments.

Fig. 8. The different non jumping transformations used in the algorithm. Each of these transformations first randomly selects an object in the current configuration of rectangles and segments, and applies a specific parameter perturbation on the selected object, except in the case of a connection perturbation (see 8(h)) where two objects are modified.

*3) Simulated annealing:* To find a minimizer of the energy $U(.)$ we use a simulated annealing framework. Instead of generating samples of $h(.)$, we simulate $h^{\frac{1}{T_t}}(.)$. The temperature parameter $T_t$ tends to zero as $t$ tends to $\infty$. Note that it is equivalent to the notation

$$f_t(\mathbf{z}) = Z_{T_t}^{-1} \exp\left(-\frac{U(\mathbf{z})}{T_t}\right). \qquad (17)$$

This technique has been widely used in image processing (see [32] for instance). If $T_t$ decreases with a logarithmic rate, then $X_t$ tends to one of the global maximizers of $h(.)$. Of course, in practice it is not possible to use a logarithmic evolution law and we eventually use a geometrical one. This last point makes the quality of the proposition kernels an important issue. As a consequence we design kernels such that the trajectory of the Markov chain is poorly correlated to insure a good exploration of the state space.

*C. Specific Transformations*

We detail in this section the proposition kernels used in our case.

*1) Birth or death transformations:* We consider two simple birth or death transformations, $Q_{BDR}$ and $Q_{BDS}$ corresponding to the random birth or death of a rectangle (BDR) or a segment (BDS). The birth update follows a uniform distribution $|.|/|S^r|$ or $|.|/|S^s|$ depending on the type of object to be created. The death update proposes removing an object uniformly selected among the current objects. The Green ratios associated with the birth and the death of a rectangle are respectively $R_{BDR}(\mathbf{z}, \mathbf{z} \cup u) = |S^r| h(\mathbf{z} \cup u)/(n(\mathbf{z})+1)h(\mathbf{z})$, and $R_{BDR}(\mathbf{z}, \mathbf{z} \setminus u) = n(\mathbf{z})h(\mathbf{z} \setminus u)/|S^r|h(\mathbf{z} \setminus u)h(\mathbf{z})$. The ratios in the case of birth or death of a segment are similar.

*2) Translations, rotations, dilatations:* We have implemented the transformations depicted in Figure 8. Each of these transformations uses a parameter $z$ that is randomly chosen in some symmetric set $\Sigma$. For instance, the rotation perturbations use a random parameter $\xi \in \Theta = [-\Delta\varphi, \Delta\varphi]$ to generate the new angle for the selected object. If $u$ is chosen uniformly in $\mathbf{z}$ and the distribution of $\xi$ is symmetric, the suitable Green ratio is given by the usual Metropolis-Hastings ratio $R(\mathbf{z}, \mathbf{z}') = h(\mathbf{z}')/h(\mathbf{z})$.

*3) Birth or death of an aligned rectangle:* This kernel proposes either to create (with probability $p_b$) or to remove (with probability $p_d = 1 - p_b$) an interacting pair of rectangles.

*a) Birth:* The birth update first selects a point $u$ of $\mathbf{z} \cap \gamma_1^r$, then generates a new point $v$ aligned with $u$ in the sense of one of the alignment relations $\sim^a l$, and proposes $\mathbf{z}' = \mathbf{z} \cup v$.

*b) Death:* The death update selects a pair of aligned points provided that at least one of them is in $\gamma_1^r$, chooses an object $v$ in this pair with probability 0.5, and proposes to remove $v$: $\mathbf{z}' = \mathbf{z} \setminus v$.

The expression for the Green ratio associated with this kind of transformations is detailed in [21] and [6], and we refer the reader to these papers for detailed explanations on how the ratios are derived.

*4) Connection perturbation:* This intuitive transformation allows to modify the segment network, by acting simultaneously on two connected segments. The different steps are 1. choose a couple of connected segments, 2. generate a random perturbation vector, 3. apply this vector to each of the connected extremities, 4. test if the two new segments are still connected, if not stop here, otherwise 5. propose $\mathbf{y}'$ obtained by replacing the two former segments by the new perturbed one. This transformation is illustrated by Figure 8(h). The Green ratio is given by $h(\mathbf{z}')/h(\mathbf{z})$.

*D. Reference measure*

For computational convenience, the reference intensity measure $\nu(.)$ usually used is uniform (see [20]). The advantage of using a simple intensity measure is that it makes the birth of a point easier. However, in our setup, points of interest are those in $\gamma_1$ which is of small Lebesgue measure. To improve the exploration of $\gamma_1$, a solution is to use a reference measure favoring this set. In [29] we show how to tune the reference measure in order to improve the mixing ability of the Markov Chain.

*E. Convergence of the algorithm*

The convergence of the algorithm holds. We derived sufficient conditions in [21] and showed in [6] that these conditions are fulfilled in the specific case of a process of rectangles. The generalization to the case of processes of both segments and rectangles is straightforward, due to the representation of the new mark space as a product. The convergence of the simulated annealing towards a global maximum of the density $h(.)$ has been proved in [22] using Dobrushin conditions.

## V. RESULTS

We present in this section a set of results on real data. We first detail the parameters of the internal field, as these parameters were tuned once and kept constant for all applied results. We then present results on a very crude DEM. This example illustrates how interesting it is to include two different types of elements. We first present a result obtained using only as process of segments, then a result obtained using only a process of rectangles and finally a result obtained using both processes. We then present results on different types of data (optical and LASER), showing the generality of the proposed model.

### A. Fixed parameters

In Table I we present the parameters defining the internal model and the space parameters. These parameters are fixed for all the results presented later in this paper. These parameters have been tuned by hand, but a supervised learning is possible, if some examples are provided (see [25]).

| Objects space | |
|---|---|
| $l^r_{\min}$ | 5m |
| $l^r_{\max}$ | 30m |
| $L^r_{\min}$ | 5m |
| $L^r_{\max}$ | 40m |
| $L^s_{\min}$ | 9m |
| $L^s_{\max}$ | 30m |

| Cooperation | |
|---|---|
| Definition | |
| $\delta d_{\max}$ | 6m |
| $\delta\theta^{rs}_{\max}$ | 30° |
| Weights | |
| $a^{\sim coop.}$ | 0.1 |
| $b^{\sim coop.}$ | 1 |

| Rectangle | |
|---|---|
| Definition | |
| $\delta C_{\max}$ | 6m |
| $\delta\theta^r_{\max}$ | 30° |
| Alignments | |
| $a^{\sim al}$ | 0.3 |
| $b^{\sim al}$ | 0.4 |
| Paving | |
| $a^{\sim pav}$ | 0.3 |
| $b^{\sim pav}$ | 0.4 |
| Exclusion | |
| $a^{\sim excl}$ | -10 |

| Segment | |
|---|---|
| Definition | |
| $\delta E_{\max}$ | $5m$ |
| $\delta\theta^s_{\max}$ | 30° |
| Flat connection | |
| $a^{\sim conn.al.}$ | 1.5 |
| $b^{\sim conn.al.}$ | 1.5 |
| Orth. connection | |
| $a^{\sim conn.orth.}$ | 0 |
| $b^{\sim conn.orth}$ | 2 |
| Exclusion | |
| $a^{\sim excl.s}$ | -15 |

TABLE I
PARAMETERS DESCRIBING THE INTERNAL FIELD OF BOTH PROCESSES AS WELL AS THE INTERACTION TERM BETWEEN RECTANGLES AND SEGMENTS. THESE PARAMETERS ARE FIXED FOR ALL THE RESULTS PRESENTED IN THIS PAPER.

### B. Crude DEM

We present in Figure 9(a) a very crude Digital Elevation Model. This DEM comes from a simulation and has been provided by the French National Geographic Institute (IGN). We consider this DEM to be very crude as it exhibits two main characteristics. First, due to the way the data have been generated, the vertical resolution of the DEM is very low, as shown by the low number of gray levels of Figure 9(a). Second, there are some occlusions: large amounts of data represented by black areas are missing.

*1) Rectangles process:* We present in Table II the data term parameters employed. In order to deal with missing data, we add the condition that a rectangle can be attractive only if the amount of data available is large enough (at least 70% of the rectangle surface). We present on Figure 9(b) a detection result obtained by using only the rectangle process. The homogeneity data term clearly plays its role. The segmentation obtained

indeed follows the homogenous areas. However, the rectangles fail to follow the discontinuities.

*2) Segments process:* We present the segment data term parameters in Table II. These parameters were tuned to deal with smooth discontinuities. On Figure 9(c), we present an extraction result obtained using the segment process only. Note that the extraction process tends to give curved linear networks.

*3) Cooperation between segments and rectangles:* We present now the extraction result obtained using the two processes together in two steps. In Figure 10(b), we show the obtained segment configuration, while on Figure 10(a) we show the rectangle configuration. The cooperation term plays its role, as both the rectangles and the segments processes fit the data better when used together. This example on a crude DEM is however disappointing. The data are indeed too crude to obtain a satisfactory extraction result, although it illustrates well the originality of our approach. As a consequence, we present in the rest of this section results on different Digital Elevation Models.

### C. Satellite DEM

We present on Figure 11(a) another simulation of a satellite DEM also provided by the IGN. This DEM is far better due its improved vertical and horizontal resolution and the rectangle extraction is therefore more meaningful. We present in Table II the data term parameters employed. Figures 11(b) and 11(c) present the extraction results obtained using both processes. These results show that our approach is interesting as it provides a kind of land register useful for further analysis. In particular, our approach provides a starting point for precise building reconstruction.

### D. Aerial DEM

We present in Figure 12(a) a DEM of a part of Rennes, France. It is a DEM obtained by stereovision on aerial images. This DEM is very noisy with respect to other data available (see [6]). The data term parameters employed are detailed in Table II. Figures 12(b) and 12(c) present the extraction result obtained using both processes. Again, the obtained result allows further 3D roof extraction.

### E. Laser DEM

We present in Figure 13(a) a DEM obtained by a Laser measurement. This piece of data originally consists in a set of sparse 3D points. The obtained DEM is consequently very smooth. In Table II we show the data term parameters employed. Figures 13(b) and 13(c) present the extraction results obtained using both processes. The results are relevant, but somehow disappointing. It seems that our approach is better suited to low quality DEMs than to precise data.

### F. Comments

The algorithm requires a large number of iterations (on average 25.000.000) and a high computational time. Each simulation takes around 6 hours on an image of size 1000

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE

(a) Crude Digital Elevation Model ©IGN

(b) Extraction result obtained using only the process of rectangle.

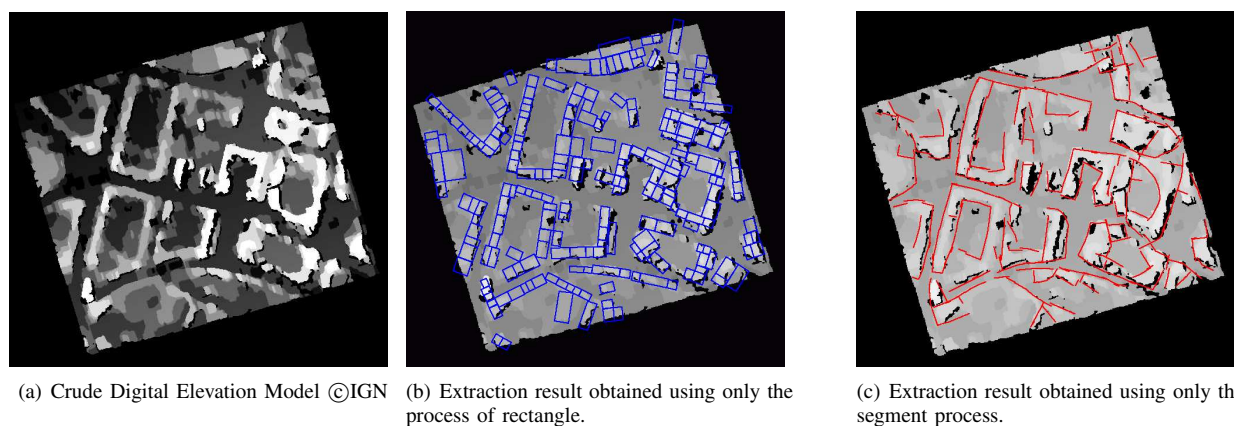(c) Extraction result obtained using only the segment process.

Fig. 9. Crude Digital Elevation Model provided by the French National Geographic Institute. The analysis of this DEM faces two major issues: first, large areas of data are missing (black areas); second, the vertical resolution is poor, limiting the amount of information available.
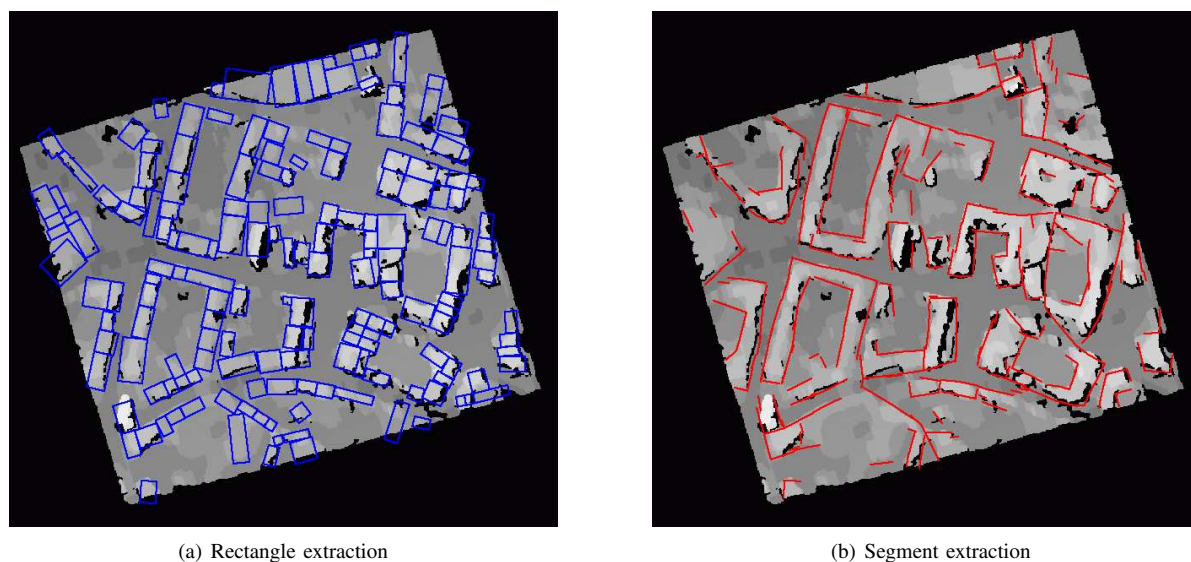


(a) Rectangle extraction

(b) Segment extraction

Fig. 10. Extraction result obtained using both processes simultaneously.
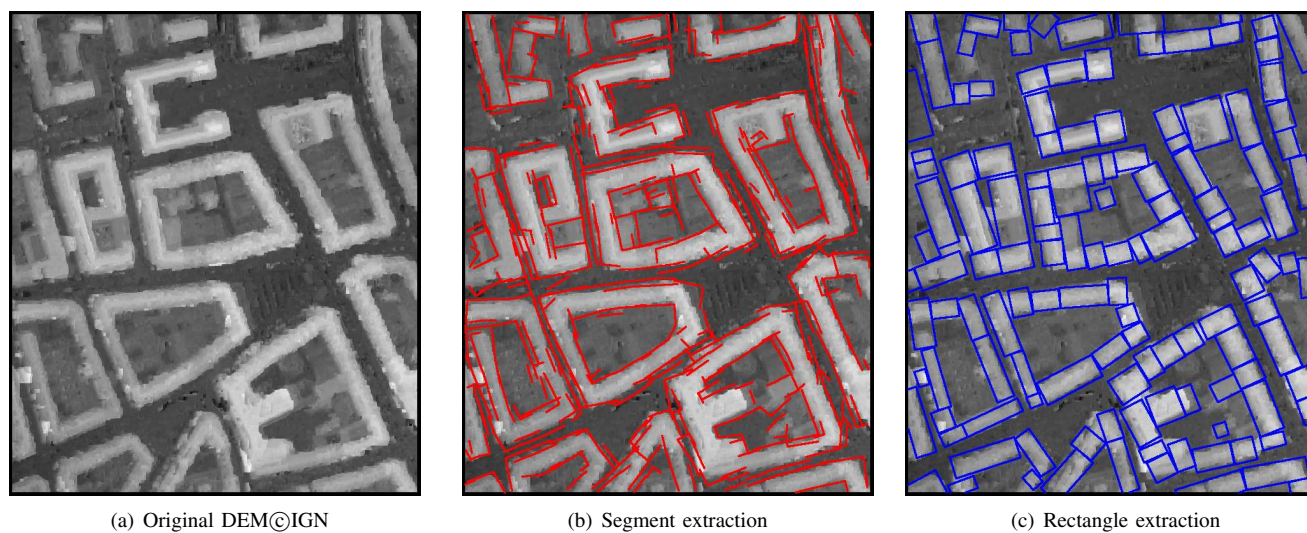


(a) Original DEM©IGN

(b) Segment extraction

(c) Rectangle extraction

Fig. 11. Digital Elevation Model of a part of Amiens, France obtained by satellite stereovision and extraction results

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE

13

| *Data term parameters associated with Figures 9 and 10* | | | | | | |
|---|---|---|---|---|---|---|
| Rectangle data term parameters | Inside points resolution<br>Distance of lateral bands | 2 m<br>10 m | $H_{\min}$<br>$v_{\min}$ | 5 m<br>90% | $\sigma_{\max}$<br>$\rho^r$ | 5 m<br>1.2 |
| Segment data term parameters | Slices discretization step<br>Distance between slices<br>Length of slices | 1 m<br>2 m<br>30 m | $\nabla H_{\min}$<br>$\delta H_{\min}$<br>$l_{sel}$ | 0.7<br>5<br>4 m | $\delta_r$<br>$\eta$<br>$\rho^s$ | 8 m<br>90%<br>3 |
| *Data term parameters associated with Figure 11* | | | | | | |
| Rectangle data term parameters | Inside points resolution<br>Distance of lateral bands | 2 m<br>10 m | $H_{\min}$<br>$v_{\min}$ | 5 m<br>90% | $\sigma_{\max}$<br>$\rho^r$ | 10 m<br>1.2 |
| Segment data term parameters | Slices discretization step<br>Distance between slices<br>Length of slices | 1 m<br>1.5 m<br>20 m | $\nabla H_{\min}$<br>$\delta H_{\min}$<br>$l_{sel}$ | 1<br>4 m<br>4 m | $\delta_r$<br>$\eta$<br>$\rho^s$ | 3m<br>90%<br>3 |
| *Data term parameters associated with Figure 12* | | | | | | |
| Rectangle data term parameters | Inside points resolution<br>Distance of lateral bands | 1m<br>10m | $H_{\min}$<br>$v_{\min}$ | 3m<br>90% | $\sigma_{\max}$<br>$\rho^r$ | 10m<br>1.2 |
| Segment data term parameters | Slices discretization step<br>Distance between slices<br>Length of slices | 0.8m<br>0.7m<br>20m | $\nabla H_{\min}$<br>$\delta H_{\min}$<br>$l_{sel}$ | 1<br>2m<br>4m | $\delta_r$<br>$\eta$<br>$\rho^s$ | 1.5m<br>90%<br>3 |
| *Data term parameters associated with Figure 13* | | | | | | |
| Rectangle data term parameters | Inside points resolution<br>Distance of lateral bands | 1 m<br>10 m | $H_{\min}$<br>$v_{\min}$ | 3 m<br>90% | $\sigma_{\max}$<br>$\rho^r$ | 10m<br>1.2 |
| Segment data term parameters | Slices discretization step<br>Distance between slices<br>Length of slices | 0.8 m<br>0.7 m<br>20 m | $\nabla H_{\min}$<br>$\delta H_{\min}$<br>$l_{sel}$ | 1<br>2 m<br>4 m | $\delta_r$<br>$\eta$<br>$\rho^s$ | 1.5m<br>90%<br>3 |

TABLE II
DATA TERM PARAMETERS EMPLOYED FOR THE DIFFERENT RESULTS PRESENTED IN FIGURES 10, 11,12 AND 13.

by 1000, including approximatively 150 buildings[2]. Improving the speed of the algorithm is therefore a major issue, although the computational time depends more on the complexity of the urban area than on the size of the image. An interesting idea would be to pre-compute the result of the discontinuity filter, since it appears that the computation associated with the segment process take most of the time. Another idea could be to keep the best objects found so far in memory.

## VI. CONCLUSION AND FUTURE WORK

We have proposed an original approach based on stochastic geometry amenable to the introduction of a prior knowledge on both the shape of the primitives to be extracted and their spatial patterns in term of their interactions.

More specifically, we have presented a model based on two types of geometrical objects. We have proposed a point process model of segments to detect discontinuities and a model of rectangles for segmenting homogeneous areas. The prior term we employ favors the connections between segments, a paving behavior of rectangles and makes both types of objects interact.

Although our approach is based on very simple objects, it proves to be powerful when applied on real data. We have indeed been able to process Digital Elevation Models of various type (from aerial/satellite stereovision data and Laser measurement). To our best knowledge, few automatic methods are able to process such a variety of data.

Future work should involve the introduction of more complex primitives (e.g. corners, roof edges, etc...). However two major issues need to be solved in order to fully exploit this kind of models. First, the learning of parameters should be carefully examined, even if the prior model parameters proved to be very robust in practice. Second, the algorithm employed is very slow. There is a huge need for proposing new algorithms to speed up the computation. A first direction is to improve the simulated annealing. Adaptive cooling schedules are a possibility we will examine in a near future. Another interesting idea would be to test whether adding a memory could improve the algorithmic performances.

## REFERENCES

[1] 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*, vol. 72, no. 2, pp. 185–203, 1998.
[2] H. Mayer, "Automatic object extraction from aerial imagery-a survey focusing on buildings," *Computer Vision and Image Understanding*, vol. 74, no. 2, pp. 138–149, 1999.
[3] 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 haute résolution," Ph.D. dissertation, ENST, Paris, France, 2002.
[4] 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.
[5] M. Fradkin, M. Roux, H. Maître, and U. Leloglu, "Surface reconstruction from multiple aerial images in dense urban areas." in *Proc of IEEE Int. Conf. on Computer Vision and Pattern Recognition*, vol. 1, Fort Collins, Colorado, USA, June 1999, pp. 262–267.
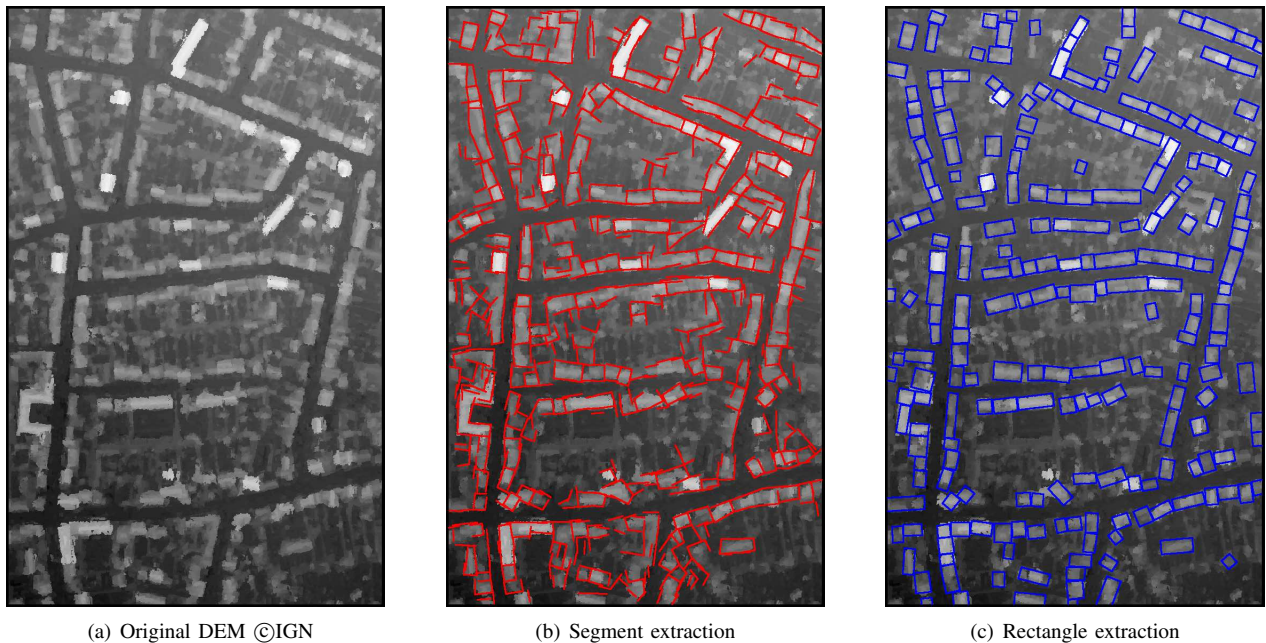
[2]Using a 3 GHz Pentium 4 machine

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE

14

(a) Original DEM ©IGN  (b) Segment extraction  (c) Rectangle extraction

Fig. 12. Digital Elevation Model of a part of Rennes, France and extraction result. This DEM has been constructed by aerial stereovision.



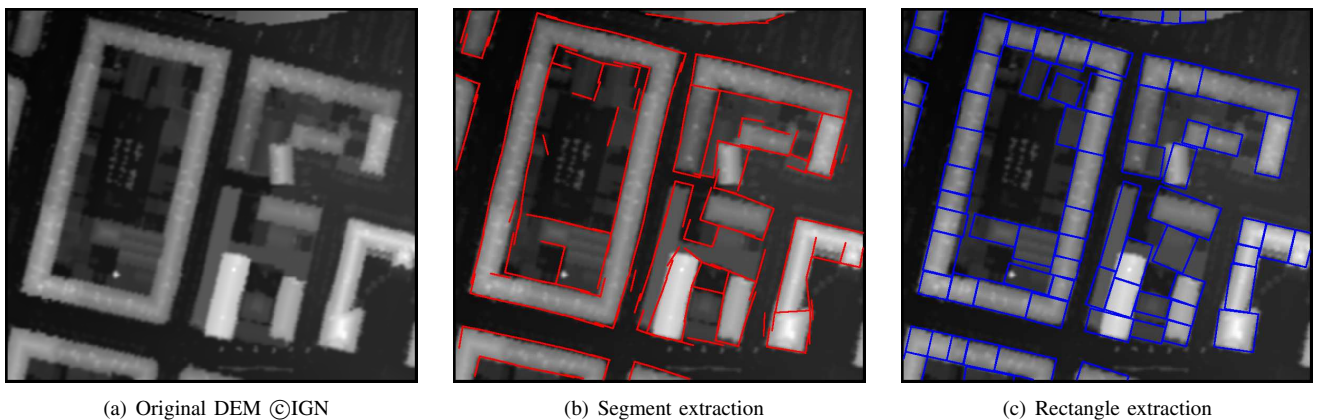(a) Original DEM ©IGN  (b) Segment extraction  (c) Rectangle extraction

Fig. 13. Digital Elevation Model of a part of Amiens, France and obtained extraction. This DEM has been obtained by Laser measurements, and exhibits smooth discontinuities.

[6] M. Ortner, X. Descombes, and J. Zerubia, "Building outline extraction from Digital Elevation Models using marked point processes," *International Journal of Computer Vision*, 2006, to appear.

[7] ——, "Point processes of segments and rectangles for building extraction from digital elevation models," in *ICASSP06*, Toulouse, France, May 2006.

[8] J. Besag, "On the statistical analysis of dirty pictures," *Journal of Royal Statistic Society*, vol. B, no. 68, pp. 259–302, 1986.

[9] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images," *IEEE Trans. on PAMI*, vol. 6, no. 6, pp. 721–741, Nov. 1984.

[10] O. Viveros-Cancino, "Analyse du milieu urbain par une approche de fusion de données satellitaires optiques et radar," Ph.D. dissertation, University of Nice-Sophia Antipolis, 2003.

[11] A. Baddeley and M. N. M. Van Lieshout, "Stochastic geometry models in high-level vision," *Statistics and Images*, vol. 1, pp. 233–258, 1993.

[12] H. Rue and M. Hurn, "Bayesian object identification," *Biometrika*, vol. 3, pp. 649–660, 1999.

[13] H. Rue and A. R. Syverseen, "Bayesian object recognition with Baddeley's delta loss," *Adv. Appl. Prob*, vol. 30, pp. 64–84, 1998.

[14] A. Pievatolo and P. Green, "Boundary detection through dynamic polygons," *Journal of the Royal Statistical Society*, vol. B, no. 60, pp. 609–626, 1998.

[15] R. Stoica, X. Descombes, and J. Zerubia, "A Gibbs point process for road extraction from remotely sensed images," *International Journal of Computer Vision*, vol. 37, no. 2, pp. 121–136, 2004.

[16] C. Lacoste, X. Descombes, and J. Zerubia, "A comparative study of point processes for line network extraction in remote sensing," *INRIA Research Report 4516*, 2002.

[17] M. Ortner, "Processus ponctuels marqués pour l'extraction automatique de caricatures de bâtiments à partir de modèles numériques d'elévation," Ph.D. dissertation, University of Nice Sophia Antipolis, 2004. [Online]. Available: http://www.inria.fr/rrrt/tu-1138.html

[18] C. Lacoste, "Extraction de réseaux liniques à partir d'images satellitaires et aériennes par processus ponctuels marqués," Ph.D. dissertation, University of Nice Sophia Antipolis, 2004, (in French).

[19] A. Srivastava, U. Grenander, G. Jensen, and M. Miller, "Jump-diffusion Markov processes on orthogonal groups for object recognition," *Journal of Statistical Planning and Inference*, 1999.

[20] M. N. M. Van Lieshout, *Markov Point Processes and their Applications*. Imperial College Press, London, 2000.

[21] M. Ortner, X. Descombes, and J. Zerubia, "A Reversible Jump MCMC sampler for building detection in image processing." in *Monte Carlo methods and quasi-Monte Carlo methods*, Juan les Pins (France), June 2004, springer Verlag.

[22] M. N. M. Van Lieshout, "Stochastic annealing for nearest-neighbour

point processes with application to object recognition." *Advances in Applied Probability*, vol. 26, pp. 281–300, 1994.

[23] D. J. Strauss, "A model for clustering," *Biometrika*, vol. 62, pp. 467–475, 1975.

[24] M. Ortner, X. Descombes, and J. Zerubia, "Un nouveau modèle pour l'extraction de caricatures de bâtiments sur des Modèles Numériques d'Elévation." in *TAIMA*, Hammamet (Tunisia), Octobre 2003, in French.

[25] C. J. Geyer, "Likehood inference for spatial point processes," in *Stochastic Geometry Likehood and computation*, O. Banorff-Nielsen, W. Kendall, and M. V. Lieshout, Eds. Chapman and Hall, 1999.

[26] M. N. M. Van Lieshout, "Markovianity in space and time," *Dynamics and stochastics: Festschrift in honour of Michael Keane*, pp. 154–167, 2006.

[27] 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.

[28] R. Stoica, "Processus ponctuels pour l'extraction des réseaux linéiques dans les images satellitaires et aériennes." Ph.D. dissertation, University of Nice Sophia-Antipolis, 2001.

[29] M. Ortner, X. Descombes, and J. Zerubia, "A marked point process of rectangles and segments for automatic analysis of Digital Elevation Models." *INRIA Research Report 5712*, October 2005.

[30] P. Green, "Reversible jump Markov chain Monte-Carlo computation and Bayesian model determination," *Biometrika*, vol. 57, pp. 97–109, 1995.

[31] C. Geyer and J. Møller, "Simulation and likehood inference for spatial point processes," *Scandinavian Journal of Statistics*, vol. Series B, 21, pp. 359–373, 1994.

[32] G. Winkler, *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: a Mathematical Introduction*. Springer-Verlag, 2003, second Edition.