INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Building Detection by Markov Object processes and a MCMC algorithm

Laurent Garcin — Xavier Descombes — Josiane Zerubia — Hervé Le Men

N° 4206

Juin 2001

THÈME 3

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4206--FR+ENG

Building Detection by Markov Object processes and a MCMC algorithm

Laurent Garcin* , Xavier Descombes , Josiane Zerubia , Hervé Le Men†

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Ariana

Rapport de recherche n° 4206 — Juin 2001 — 31 pages

Abstract: This work aims at detecting buildings in digital aerial photographs. Here we model a set of buildings by a configuration of objects. We define a point process on the set of configurations, which splits into two parts :

- the first one is a prior model on the configurations which use interactions between objects,
- the second one is a data model which enforces the coherence with the image.

Thus we have a posterior distribution whose maximum has to be found. In order to achieve this maximum, we use a MCMC simulation - a Metropolis-Hastings-Green algorithm - mixed with a simulated annealing. Then we test this method on both synthetic and real stereo-images.

Key-words: Point processes, RJMCMC, stereovision, building extraction

* INRIA et Institut Géographique National
† Institut Géographique National, St Mandé

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, B.P. 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : 04 92 38 77 77 - International : +33 4 92 38 77 77 — Fax : 04 92 38 77 65 - International : +33 4 92 38 77 65

Détection de bâtiments par un processus de Markov objet et un algorithme de type MCMC

Résumé : Le but de ce travail est de détecter les bâtiments à partir de photographies aeriennes numériques. Nous modélisons un ensemble de bâtiments par une configuration d'objets. Nous définissons un processus ponctuel sur l'ensemble des configurations qui se décompose en deux parties :

- la première est un modèle a priori sur les configurations qui considère des interactions entre les objets,
- la seconde est un modèle d'attache aux données qui induit la cohérence du résultat avec l'image traitée.

Nous avons ainsi une distribution a posteriori dont nous recherchons la configuration maximale. Pour obtenir ce maximum, nous utilisons une simulation de type MCMC - un algorithme de Metropolis-Hasting-Green- couplée avec un schéma de recuit simulé. Nous testons la méthode décrite à la fois sur des données synthétiques et des images stéréoscopiques réelles.

Mots-clés : Processus ponctuels, RJMCMC, stéréovision, extraction de bâtiments

Contents

1 Introduction

Automated techniques for building reconstruction play a leading role in a wide range of domains such as telecommunication or urban planning, which demand both precise and up to date 3D cartographic information about buildings. A classical approach of this kind of problems can be found in [18], for example.

In this work we use probabilistic methods. Probabilistic techniques in image segmentation generally adopt a pixel approach. Yet the increasing resolution of the observed images leads us to take into account the geometrical information contained in images. Thus, instead of dealing with pixels, we define geometrical objects related to the image (see section 3). This is what stochastic geometry is about [1]. In order to use this concept of "object", we will need some notions about point processes (see section 4). Then we build Markov processes on the geometrical objects [11], in the same way that one builds Markov random fields on pixels. Hence we are able to design interactions between objects at a larger range than by a classical pixel approach. While a standard building reconstruction deals with each building separately, the proposed approach benefits from the prior knowledge we have about the relations between buildings (e.g. alignments). That is why we use those Markov object processes for building detection in dense urban areas. By doing so, we expect the information lost due to occlusion or shading to be replaced by the prior knowledge about the relations between buildings.

Then a problem arises which is the simulation of such processes. As some progress has recently been made in this field, we have decided to simulate our process with a Monte Carlo Markov Chain method (Reversible Jump MCMC [5]).

2 Overview

We call a "configuration" a set of objects $x = \{u_1, \cdots u_n\}$ of variable size (this will be defined more accurately in section 4). In the present case, those objects are buildings.

Then we define a distribution π on the configuration set, which satisfies some given criteria related to the prior model (see section 5) and to the data model (see section 6).

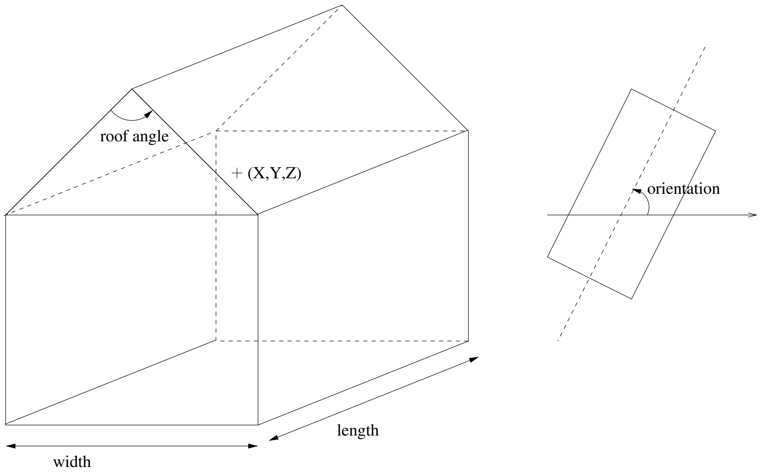


Figure 1: Building model

3 Geometrical Object Model

As we said formerly an object stands for a building. First we have to design the geometrical model of an object in order to be able to define the configuration model. Thus we represent a building by its roof (see Fig. 1).

The object parameters are :

- the 3D coordinates of the center of the roof (X, Y, Z)
- the width and length of the building (l, L)
- the angle defined by the two sides of the roof α
- the orientation of the building θ

Remark : the parameters define a 3D model so that we can use stereoscopy.

4 Point Process

We present briefly the notion of point process. Indeed, this kind of process is adapated to the concept of configuration introduced in section 2, because it allows us to define a measure on a set of n-tuples, i.e. the configuration set, more precisely defined now. However one can read [2] for further details.

4.1 Definition of a point process

Let U be a bounded Borel set of $\mathbb{R}^d$, $\mathcal{B}$ the Borel algebra of U and λ the Lebesgue measure on $\mathbb{R}^d$.

Remark : from our point of view, U is the set in which the object parameters take their values.

Then we define a measured space $(\Omega, \mathcal{F}, \mu)$ by the following way :

$$\Omega = \bigcup_{n=0}^{\infty} \Omega_n, \quad \Omega_n = \{\{x_1, \cdots, x_n\} \subset U\} \tag{1}$$

Ω is the configuration space. We denote $\mathcal{F}$ the σ-algebra associated to Ω and for $F \in \mathcal{F}$:

$$\mu(F) = e^{-\lambda(U)} \left(\mathbb{1}(\emptyset \in F) + \sum_{n=1}^{\infty} \frac{1}{n!} \int \cdots \int \mathbb{1}(\{x_1, \cdots, x_n\} \in F) \lambda(dx_1) \cdots \lambda(dx_n) \right) \tag{2}$$

We call "point process" a measurable application from a probabilist set to $(\Omega, \mathcal{F}, \mu)$. Now we are allowed to deal with the density of a "point process" with respect to the measure μ.

So the distribution π of section 2 can be written as : $\pi(dx) = h(x)\mu(dx)$. The density h splits into two parts $h(x) = f(x)g(x)$, where f stands for the prior model and g stands for the data model.

4.2 An example of point process : Markov process

Let $\sim$ be a symmetric relation on U. Then we define the neighbourhood of $A \subset U$ by the following way :

$$V(A) = \{u \in U | u \sim v, v \in A, u \notin A\} \tag{3}$$

A set $x = \{x_1, \cdots, x_n\}$ is called a clique if all the points of x are neighbours.

Remark : we also give the definition of the subset of the neighbours of a set $A \subset U$:

$$\mathcal{V}(A) = \{u \in A | u \sim v, v \in A\} \tag{4}$$

We say that $f : \Omega \rightarrow [0, +\infty)$ is a Markov function with respect to the relation $\sim$ if for $u \in \Omega$:

- $f(x) > 0 \Rightarrow f(y) > 0, \forall y \subset x$

- if $f(x) > 0$ and $u \in U$ then : $\rho(x,u) = \frac{f(x \cup u)}{f(x)}$ only depends on u and on $V(u) \cap x$. $\rho(x,u)$ is the so called Papangelou conditional intensity.

A Markov point process is a point process whose density is a Markov function.

A function $g : \Omega \rightarrow [0, +\infty)$ is called “interaction function” if $g(x) \neq 1$ implies that x is a clique. Then we have the following equivalence : f is a Markov function if and only if there exists an interaction function g such that, for $x \in \Omega$:

$$f(x) = \prod_{y \subset x} g(y) \tag{5}$$

This result is similar to the Hammersley-Clifford theorem for the Markov random fields. In fact, as for Markov random fields, we can define clique potentials and express the density f through these potentials :

$$f(x) \propto exp\left(-\sum_{c \in \mathcal{C}(x)} U(c)\right) \tag{6}$$

where $\mathcal{C}(x)$ represents the set of all the cliques in x.

If we call $\mathcal{C}_r(x)$ the set of the cliques with r elements, then we can also write f as follows :

$$f(x) \propto exp\left(-\sum_{r} \sum_{c \in \mathcal{C}_r(x)} U_r(c)\right) \tag{7}$$

where U_r is a potential defined on the cliques of r elements.

5 Prior Model

We choose the density of the prior model as follows :

$$f(x) \propto \beta^{n(x)} \alpha(x) \tag{8}$$

- $n(x)$ is the number of objects in the configuration x,
- the parameter β controls the intensity of the point process,
- the function α models the interactions between the objects of the configuration.

We only use pairwise interaction (see further).

$$\alpha(x) = \prod_{\substack{\{u,v\} \subset x \\ u \sim v}} \varphi(u,v) \tag{9}$$

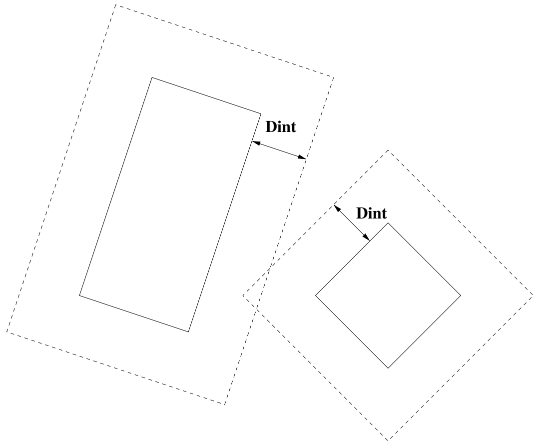


Figure 2: Relation $\sim$

5.1 Relations

We define two types of relations :

1. Relation $\sim$: We say that two objects are in relation $\sim$ when the frames around them intersect (see Fig. 2).

2. Relation $\simeq$: If we call C_1, C_2 the centers of two buildings, L_1, L_2 their respective lengths, ϕ and ψ defined as in the Fig. 3. We say that the two buildings are neighbours with respect to the relation $\simeq$ if :

$$\begin{aligned} 0 < d(C_1, C_2) - (L_1 + L_2)/2 < d_{max} \\ |\phi| < \Delta\theta_{max}, \quad |\psi| < \Delta\theta_{max} \end{aligned} \tag{10}$$

   This relation is designed for buildings neighbouring along their width so we define a similar relation for buildings neighbouring along their length.

5.2 Interactions

There are five kinds of pairwise interaction :

- for the relation $\sim$:

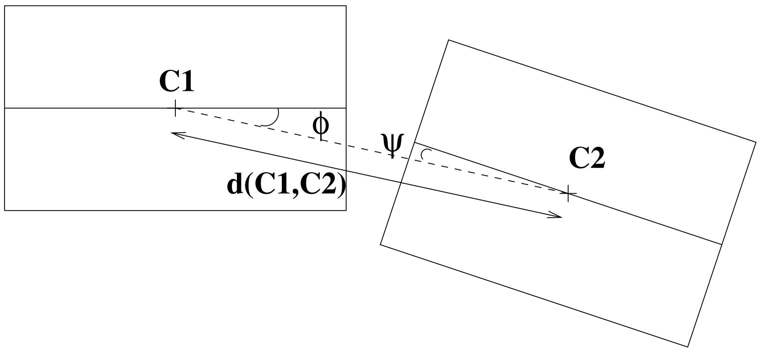


Figure 3: Relation $\simeq$

  - a term favours identical orientations (see Fig. 4).
  - a term penalizes intersections.

- for the relation $\simeq$:
  - a term favours the closeness of the buildings centers (see Fig. 5).
  - a term favours equal sizes of two buildings facing edges.
  - a term favours equal rooves angles.
  - a term favours equal rooves heights.

A potential is associated to each kind of interaction (see Fig. 6). The parameter may be the orientation difference, the intersection area, the distance between centers, the difference between lengths, widths or rooves angles.

The function ϕ is designed as follows :

$$\phi(u, v) \propto exp(-U_p(u, v)) \tag{11}$$

where $u, v \in U$ ans U_p is the sum of the energy terms related to the six different kinds of interaction.

6 Data model

6.1 Computation of the data model terms

We suppose that we know all the viewpoint parameters i.e. :

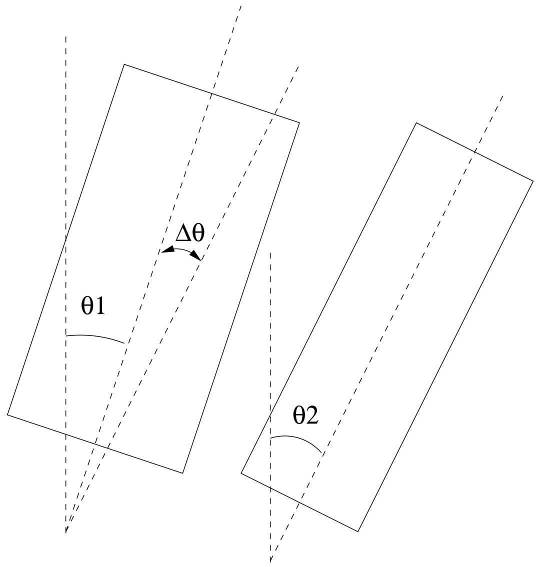


Figure 4: Orientation difference

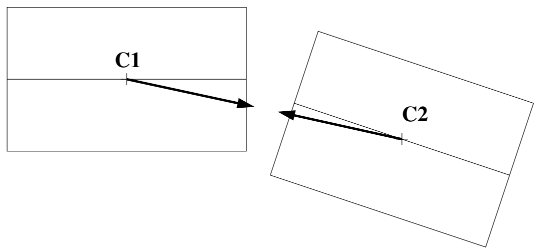


Figure 5: Bringing the centers closer

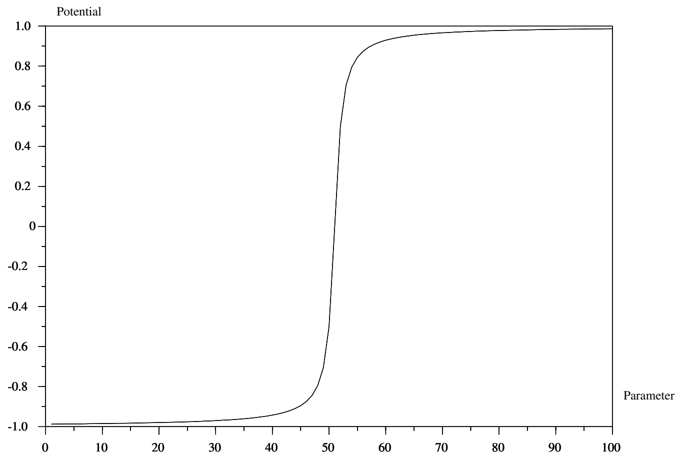


Figure 6: Potential

- the intrinsinc parameters :
  - the focal distance
  - the pixel size
- the extrinsinc parameters :
  - the camera position
  - the camera orientation

Thus we can project the configuration onto the image referential (see Fig. 7) and compute the data fitting term.

We use the following data model :

$$g(x) = \prod_{u \in x} \psi(u) \tag{12}$$

It means that the data model is computed separately for each object. We use three kinds of data fitting terms :

- the first one favours low variance on each roof side.

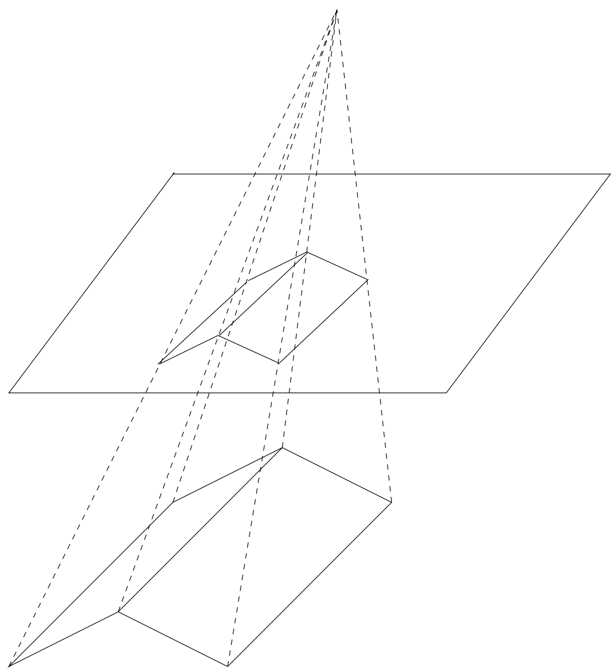


Figure 7: Projecting a roof onto the image

- the second one favours high gradient on the roof edges.
- the third one favours high contrast between the two roof sides.

We formulate again the data fitting term with an energy expression :

$$\psi(u) \propto exp(-U_a(u)) \tag{13}$$

where $u \in U$ and U_a is the sum of the three energy terms defined above.

6.2 Multiview

If we have several images, then the configuration can be projected into each image and the data fitting term can be computed for each image. By the way, this is essential if we want to build 3D objects : only stereoscopy can theoretically ensures the uniqueness of each 3D object. The object must correspond to a minimum of the data potential for each image. So this approach is different from the classical correlation-based stereoscopy. The greater the number of images, the more reliable and accurate the result.

7 Choosing the potentials

7.1 Potential types

We distinguish three types of potential (see Fig. 8) :

- Attractive potential : $U < 0$
- Repulsive potential : $U > 0$
- Mixed potential

For each kind of interaction or data fitting term, we choose a potential type :

- An attractive potential to bring closer the building centers
- A repulsive potential to penalize the buildings intersections
- A Mixed potential for the other kinds of interaction :
  - Orientation difference
  - Width or length difference

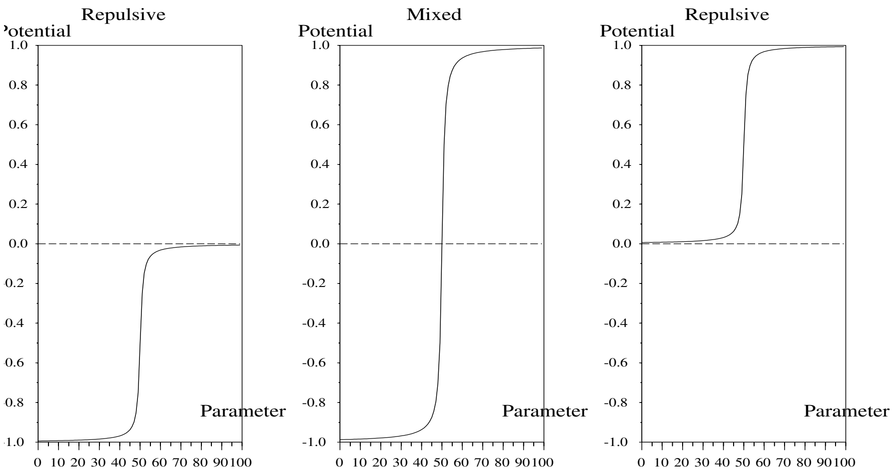


Figure 8: Potential types

  - Roof angle difference

- A Mixed potential for the data fitting terms :
  - Variance
  - Gradient
  - Contrast

The slopes chosen for the potentials play a great role (see Fig. 9) : the steeper the slope is, the more the associated potential has a "threshold" behaviour. On the contrary, a less steep slope allows a more flexible behaviour.

During the simulation, steep slopes guarantee accuracy but not exhaustivity whereas soft slopes guarantee a better exhaustivity but not accuracy. In both cases, the solution to overcome the problem of exhaustivity or accuracy is to increase the simulation time.

7.2 Energy parameters

We normalize all potentials between -1 and 0, 0 and 1 or -1 and 1 whether the potential is attractive, repulsive or mixed. Then we can express the total energy of

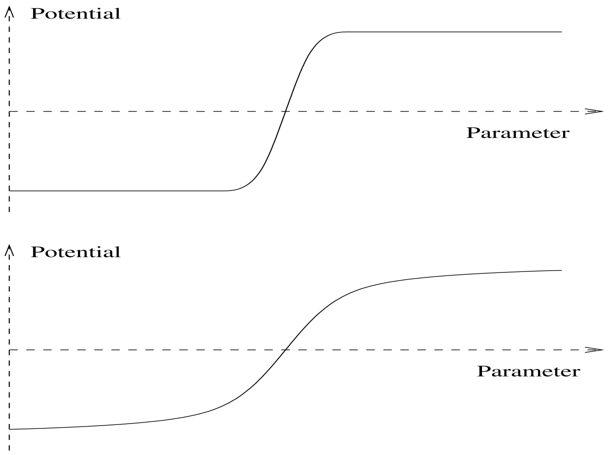


Figure 9: Potential slope

the configuration as $U(x) = \sum \gamma_i U_i(x)$ where the U_i are the normalized potentials. The parameters γ_i are chosen according to the relative importance we wish to grant to each type of interaction or data fitting term.

8 Simulation

8.1 Building a Markov chain : Reversible Jump Monte Carlo Markov Chains

The simulation consists in building a Markov chain which converges to the distribution π. So we have to define a transition kernel P which is aperiodic, irreductible and π-invariant. In fact, we replace the invariance condition by a reversibility condition which is stronger and easier to verify. The reversibility condition means that $\int_A \pi(dx)P(x,B)$ must be symmetric with respect to A and B. Indeed, by replacing B by Ω, we obtain the invariance condition.

Let $Q(x,dy)$ be a transition kernel. We define another kernel P which is defined as follows : we choose a new state y from the current state x and the transition kernel Q and we accept this proposal with a probability $r(x,y)$ or we stay at the state x otherwise. So we can write :

$$P(x,B) = \int_B Q(x,dy)r(x,y) + s(x)\mathbb{1}(x \in B) \tag{14}$$

$$s(x) = \int_\Omega Q(x,dy)(1 - r(x,y)) \tag{15}$$

Consequently, the reversibility condition can be written for all Borel sets $A, B \in \Omega$:

$$\begin{aligned} &\int_A \pi(dx) \int_B Q(x,dy)r(x,y) + \int_{A\cap B} \pi(dx)s(x) \\ = &\int_B \pi(dy) \int_A Q(y,dx)r(y,x) + \int_{B\cap A} \pi(dy)s(y) \end{aligned} \tag{16}$$

In order to verify the reversibility condition, it is sufficient to verify that for all A, B :

$$\int_A \pi(dx) \int_B Q(x,dy)r(x,y) = \int_B \pi(dy) \int_A Q(y,dx)r(y,x) \tag{17}$$

If we suppose that $\pi(dx)Q(x,dy)$ has a density with respect to a symmetric measure ξ on $\Omega \times \Omega$, we can write :

$$\int_A \pi(dx) \int_B Q(x,dy)r(x,y) = \int_A \int_B \xi(dx,dy)f(x,y)r(x,y) \tag{18}$$

$$\int_B \pi(dy) \int_A Q(y,dx)r(y,x) = \int_B \int_A \xi(dy,dx)f(y,x)r(y,x) \tag{19}$$

Hence the reversibility condition is satisfied as soon as :

$$f(x,y)r(x,y) = f(y,x)r(y,x) \tag{20}$$

We want to maximize the acceptation ratio $r(x,y)$ so we take :

$$r(x,y) = min\left(1, \frac{f(y,x)}{f(x,y)}\right) \tag{21}$$

Suppose we have defined several kernels P_m by the same way, that is we have chosen kernels Q_m and acceptance ratio r_m so that each kernel P_m verifies the reversibility condition. Obviously, $P = \sum_m p_m P_m$ ($\sum_m p_m = 1$) also verifies this condition.

Now we define the following algorithm :

- Choose a kernel Q_m according to the probabilities p_m.
- Choose a new configuration y from $Q_m(x,.)$ where x is the current configuration.
- Calculate the ratio $r_m(x,y) = \frac{f_m(y,x)}{f_m(x,y)}$.
- Accept the new configuration y with probability $min(1, r_m(x,y))$, otherwise stay at the current configuration x.

Thus the Markov chain defined by this algorithm converges to the distribution π (provided it is also aperiodic and irreductible – which is the case in practice).

8.2 Moves

Now we define what those kernels concretely represent. Indeed the kernels are associated to "moves" which will be listed further. In practice, the proposition of the kernel is made as follows :

- First uniformly choose a parameter σ in a given set $\Sigma \in \mathbb{R}^k$
- From the current configuration x and the parameter σ, propose a new configuration y.

From now on, we call Σ the auxiliary probabilistic parameter set.

Here is the list of the moves :

1. **Birth and death**

   The birth consists in adding to the current configuration an object uniformly chosen in U and the death consists in removing an uniformly chosen object from the current configuration.

2. **Translation**

   We uniformly choose a building in the current configuration and a direction among the three ones :

   - direction of the building's length
   - direction of the buliding's width
   - direction of the vertical axis

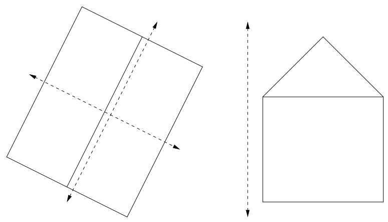


Figure 10: Translation

Then we move the roof center of the chosen building in the chosen direction of a certain distance uniformly chosen respectively on $[-\delta L_{max}; +\delta L_{max}]$, $[-\delta l_{max}; +\delta l_{max}]$ or $[-\delta z_{max}; +\delta z_{max}]$.

3. **Rotation**

   We uniformly choose a building in the current configuration and an angle in $[-\delta\theta_{max}, +\delta\theta_{max}]$. Then we rotate the chosen building with the chosen angle.

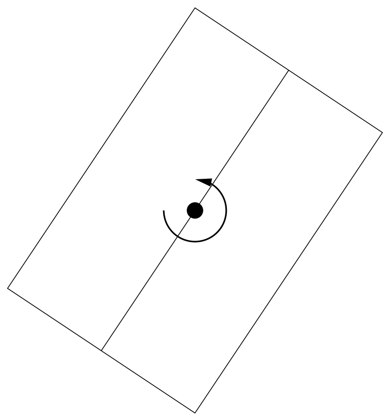


Figure 11: Rotation

4. **Dilation**

   We uniformly choose a building in the current configuration then we choose one of its four edges and we move it with a distance chosen again in a given interval.

5. **Birth and death of neighbouring buildings**

This kind of move is designed to fit the prior model. Indeed, as the neighbouring buildings are common in urban areas and as they are favoured by the prior model, it is useful to propose this kind of move since there is a good chance that it is accepted. One can refer to [15] to see other examples of prior model fitting moves for road detection.

The birth consists in uniformly choosing a building in the configuration and to add a building in its neighbourhood by the following way :

- Choose a length $L \in [L_{min}; L_{max}]$.
- Choose a distance $d \in [0; d_{max}]$.
- Choose two angles ϕ and ψ in $[-\delta\theta_{max}; +\delta\theta_{max}]$.
- Choose $\delta l \in [-\delta l_{max}; +\delta l_{max}]$.
- Choose $\alpha \in [\alpha_{min}; \alpha_{max}]$.
- Eventually, choose $Z \in [Zmin; Zmax]$.

We define the new building as shown in Fig. 12:

$$\begin{aligned} L2 &= L & l_2 &= l_1 + \delta l \\ \alpha_2 &= \alpha & \theta_2 &= \theta_1 + \phi + \psi \\ d(C_1, C_2) &= \frac{L_1 + L_2}{2} + d & Z_2 &= Z \end{aligned}$$

One can notice that the neighbouring building parameters verify the equations (10) which define the relation $\simeq$.

Thereafter we call η the application which associates a neighbouring building to another one and some auxiliary probabilistic parameters $\sigma = (L, d, \phi, \psi, \delta l, \alpha, Z)$.

$$\begin{aligned} \eta \ &: U \times \Sigma \rightarrow U \\ &(u, \sigma) \rightarrow v \end{aligned} \tag{22}$$

The death consists in removing a building uniformly chosen in the set of the neighbouring buildings of the current configuration.

Remark : in fact there are four types of neighbouring birth/death, one per building edge.

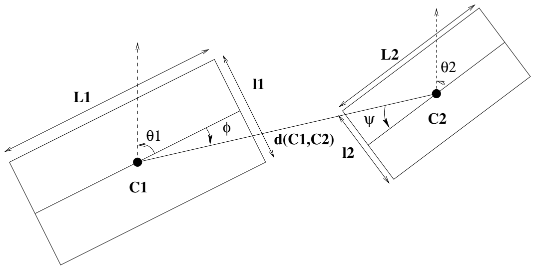


Figure 12: Birth of a new building

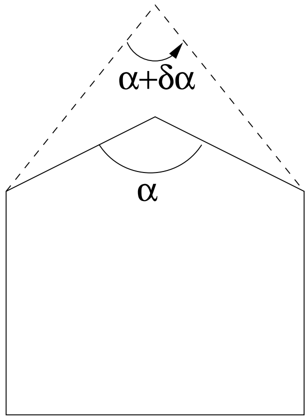


Figure 13: Roof angle modification

6. **Roof angle modification**

   We add to the roof angle a value chosen in $[-d\alpha_{max}; +d\alpha_{max}]$.

7. **Swapping length and width**

   This move is useful when a building has been detected but has a wrong orientation. The variance term favours the best of the two possibilities.

8. **Offset of a half roof**

   A half roof of a building may correspond to the wrong half roof in the image. For instance, it is the case when a building is along a street. Indeed, one of the half roof corresponds to a real half roof in the image whereas the other one is

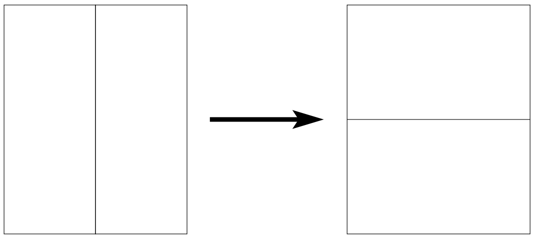


Figure 14: Swapping length and width

placed upon the street (the variance is small although there are no edges). An offset of a half roof places the building in the right position.

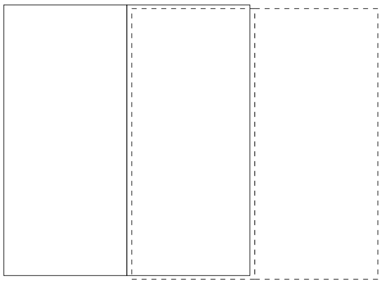


Figure 15: Offset of a half roof

General remark on the types of moves

There is a big difference between the birth/death moves on one hand and the modification of an object on the other hand. The first type of moves allows us to coarsely place some buildings whereas the second one refines the position and the shape of the coarsely placed objects. Indeed, in practice the second type of moves only propose small displacements or modifications of the objects. This kind of moves is essential because a pure death/birth process is irrealistic : the probability to propose the birth of a well placed and shaped object is extremely low.

8.3 Computation of the acceptance ratios

8.3.1 Other moves than birth and death

One can formalize these moves by the following way :

$$\begin{aligned} \zeta \ &: U \times \Sigma \rightarrow U \\ &(u, \sigma) \rightarrow v \end{aligned} \tag{23}$$

where Σ is a bounded subset of $\mathbb{R}^k$. We note ν the Lebesgue measure on $\mathbb{R}^k$.

Remark : the set Σ can be empty as in the case of the offset or inversion move.

The kernel Q corresponding to those moves is written as :

$$Q(x, A) = \sum_{u \in x} \frac{1}{n(x)} \int_{\{\sigma \in \Sigma | x \setminus u \cup \zeta(u,\sigma) \in A\}} \frac{\nu(d\sigma)}{\nu(\Sigma)} \tag{24}$$

For Borel sets $A, B \in \Omega$, let us define:

$$\xi(A \times B) = \int_A \sum_{u \in x} \int_{\{\sigma \in \Sigma) | x \setminus u \cup \zeta(u,\sigma) \in B\}} \mu(dx)\nu(d\sigma) \tag{25}$$

But, in practice, ζ is designed so that $v = \zeta(u, \sigma) \Leftrightarrow u = \zeta(v, -\sigma)$ and ξ is symmetric. Then for $x \in \Omega$, $\sigma \in \Sigma$, $u \in x$ and $y = x \setminus u \cup \zeta(u, \sigma)$:

$$\begin{aligned} \pi(dx)Q(x, dy) &= h(x)\mu(dx)Q(x, dy) && (26) \\ \xi(dx, dy) &= \mu(dx)n(x)Q(x, dy)\nu(\Sigma) && (27) \\ f(x, y) &= \frac{\pi(dx)Q(x, dy)}{\xi(dx, dy)} = \frac{1}{n(x)} \frac{h(x)}{\nu(\Sigma)} && (28) \end{aligned}$$

Consequently, according to equation (21) : $r(x, y) = min\left(1, \frac{h(y)}{h(x)}\right)$.

8.3.2 Simple birth and death

In this case, the size of the configuration varies. We have to use the *"dimension matching"* principle introduced by Green in [5]. The kernel Q proposes to add an object uniformly chosen in U with probability P_b and to remove an element with probability P_d ($P_b + P_d = 1$). Thus we can write $Q = P_b Q_b + P_d Q_d$ with :

$$\begin{aligned} Q_b(x, A) &= \int_{u \in U | x \cup u \in A} \frac{\lambda(du)}{\lambda(U)} && (29) \\ Q_d(x, A) &= \sum_{u \in x} \mathbb{1}(x \setminus u \in A) \frac{1}{n(x)} && (30) \end{aligned}$$

We define the symmetric measure ξ for Borel sets $A, B \in \Omega$:

$$\xi(A \times B) = \int_A \int_{\{u \in U | x \cup u \in B\}} \lambda(du)\mu(dx) + \int_A \sum_{u \in x} \mathbb{1}(x \setminus u \in B)\mu(dx) \tag{31}$$

Let $A_n = A \cap \Omega_n$ and $B_{n-1} = B \cap \Omega_{n-1}$ then :

$$\xi(A_n \times B_{n-1}) = \frac{e^{-\lambda(U)}}{n!} \int_{\Omega_n} \sum_{u \in U} \mathbb{1}(x \in A_n, x \setminus u \in B_{n-1})\lambda^n(dx) \tag{32}$$

$$= \frac{e^{-\lambda(U)}}{n!} \int_{\Omega_n} n\mathbb{1}(\{x_1, \cdots, x_n\} \in A_n, \{x_1, \cdots, x_{n-1}\} \in B_{n-1})\lambda^n(dx) \tag{33}$$

$$= \frac{e^{-\lambda(U)}}{(n-1)!} \int_{\Omega_{n-1}} \int_U \mathbb{1}(y \in B_{n-1}, y \cup u \in A_n)\lambda^{n-1}(dy)\lambda(du) \tag{34}$$

$$= \xi(B_{n-1} \times A_n) \tag{35}$$

That proves that ξ is symmetric.

If $y = x \cup u$ ($x \in \Omega$, $u \in U$) then :

$$\pi(dx)Q(x, dy) = h(x)\mu(dx)P_b \frac{\lambda(du)}{\lambda(U)} \tag{36}$$

$$\xi(dx, dy) = \mu(dx)\lambda(du) \tag{37}$$

$$f(x, y) = \frac{\pi(dx)Q(x, dy)}{\xi(dx, dy)} = P_b \frac{h(x)}{\lambda(U)} \tag{38}$$

And we also have :

$$\pi(dy)Q(y, dx) = h(y)\mu(dy)P_d \frac{1}{n(y)} \tag{39}$$

$$\xi(dy, dx) = \mu(dy) \tag{40}$$

$$f(y, x) = \frac{\pi(dy)Q(y, dx)}{\xi(dy, dx)} = P_d \frac{h(y)}{n(y)} \tag{41}$$

So the acceptance ratio of a birth defined by Eq. 21 is :

$$r(x, y) = min\left(1, \frac{P_d}{P_b} \frac{h(y)}{h(x)} \frac{\lambda(U)}{n(x) + 1}\right) \tag{42}$$

And for a death (move from y to x), it is the contrary.

8.3.3 Birth and death of neighbouring buildings

The kernel Q proposes to add a neighbouring building with probability P_b and to remove one of the buildings having a neighbour in the current configuration with probability P_d. We have again $Q = P_b Q_b + P_d Q_d$ with :

$$Q_b(x,A) = \sum_{u \in x} \frac{1}{n(x)} \int_{\{\sigma \in \Sigma | x \cup \eta(u,\sigma) \in A\}} \frac{\lambda(d\sigma)}{\lambda(\Sigma)} \tag{43}$$

$$Q_d(x,A) = \sum_{u \in \mathcal{V}(x)} \mathbb{1}(x \setminus u \in A) \frac{1}{nv(x)} \tag{44}$$

Let A, B be two Borel sets in Ω :

$$\xi(A \times B) = \int_A \int_{\{u \in U | u \in \eta(x,\Sigma), x \cup u \in B\}} \lambda(du)\mu(dx) + \int_A \sum_{u \in \mathcal{V}(x)} \mathbb{1}(x \setminus u \in B)\mu(dx) \tag{45}$$

We define A_n and B_{n-1} as we did previously so that :

$$\xi(A_n \times B_{n-1}) = \frac{e^{-\lambda(U)}}{n!} \int_{\Omega_n} \sum_{u \in \mathcal{V}(x)} \mathbb{1}(x \in A_n, x \setminus u \in B_{n-1}) \lambda^n(dx) \tag{46}$$

$$= \frac{e^{-\lambda(U)}}{n!} \int_{\Omega_n} \sum_{u \in x} \mathbb{1}(x \in A_n, x \setminus u \in B_{n-1}, u \in \mathcal{V}(x)) \lambda^n(dx) \tag{47}$$

$$= \frac{e^{-\lambda(U)}}{n!} \int_{\Omega_n} n \mathbb{1}(\{x_1, \cdots, x_n\} \in A_n, \{x_1, \cdots, x_{n-1}\} \in B_{n-1}) \mathbb{1}(x_n \in V(\{x_1, \cdots, x_{n-1}\})) \lambda^n(dx) \tag{48}$$

$$= \frac{e^{-\lambda(U)}}{(n-1)!} \int_{\Omega_{n-1}} \int_U \mathbb{1}(y \in B_{n-1}, u \in V(y), y \cup u \in A_n) \lambda^{n-1}(dy)\lambda(du) \tag{49}$$

$$= \xi(B_{n-1} \times A_n) \tag{50}$$

Again the measure ξ is symmetric.

Let $x \in \Omega$, $u \in x$, $\sigma \in \Sigma$, $v = \phi(u,\sigma)$ and $y = x \cup v$.

$$\pi(dx)Q(x,dy) = h(x)\mu(dx)P_b \frac{1}{n(x)} \frac{\lambda(dv)}{\lambda(\Sigma) \left| \frac{\partial \eta}{\partial \sigma} \right|} \tag{51}$$

$$\xi(dx,dy) = \mu(dx)\lambda(dv) \tag{52}$$

$$f(x,y) = \frac{\pi(dx)Q(x,dy)}{\xi(dx,dy)} = P_b \frac{h(x)}{\lambda(\Sigma) \left| \frac{\partial \eta}{\partial \sigma} \right|} \tag{53}$$

Remark : it is important to notice that $\eta(u,.) : \Sigma \rightarrow U$ is a bijection so we can write the equalities involving the jacobian term. Especially, the auxiliary parameters set Σ and the geometric parameters set have the same dimension. That is why we can speak of "dimension matching".

Similarly, we have :

$$\pi(dy)Q(y,dx) = h(y)\mu(dy)P_m\frac{1}{nv(y)} \tag{54}$$

$$\xi(dy,dx) = \mu(dy) \tag{55}$$

$$f(y,x) = \frac{\pi(dy)Q(y,dx)}{\xi(dy,dx)} = P_m\frac{h(y)}{nv(y)} \tag{56}$$

Hence the acceptation rate defined by Eq. 21 of a neighbouring birth is :

$$r(x,y) = min\left(1, \frac{P_m}{P_n}\frac{h(y)}{h(x)}\frac{\lambda(\Sigma)n(x)}{nv(x)+1}\left|\frac{\partial\eta}{\partial\sigma}\right|\right) \tag{57}$$

And for a death (move from y to x), it is the contrary.

Effective computation of the jacobian

$$X_2 = X_1 + \left(\frac{L_1+L}{2}+d\right)\cos(\theta_1+\phi) \qquad L_2 = L$$

$$Y_2 = Y_1 + \left(\frac{L_1+L}{2}+d\right)\sin(\theta_1+\phi) \qquad l_2 = l_1+\delta l$$

$$Z_2 = Z \qquad \alpha_2 = \alpha \qquad \theta_2 = \theta_1+\phi+\psi$$

$$\begin{aligned}
\left|\frac{\partial\eta}{\partial\sigma}\right| &= \begin{vmatrix}
\frac{1}{2}\cos(\theta_1+\phi) & \frac{1}{2}\sin(\theta_1+\phi) & 0 & 1 & 0 & 0 & 0 \\
\cos(\theta_1+\phi) & \sin(\theta_1+\phi) & 0 & 0 & 0 & 0 & 0 \\
-\left(\frac{L_1+L}{2}+d\right)\sin(\theta_1+\phi) & \left(\frac{L_1+L}{2}+d\right)\cos(\theta_1+\phi) & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0
\end{vmatrix} \\
&= \begin{vmatrix}
\cos(\theta_1+\phi) & \sin(\theta_1+\phi) \\
-\left(\frac{L_1+L}{2}+d\right)\sin(\theta_1+\phi) & \left(\frac{L_1+L}{2}+d\right)\cos(\theta_1+\phi)
\end{vmatrix} \\
&= \left(\frac{L_1+L}{2}+d\right) = d(C(u),C(v))
\end{aligned}$$

where $C(u)$ and $C(v)$ are the 2D centers of the objects u and v.

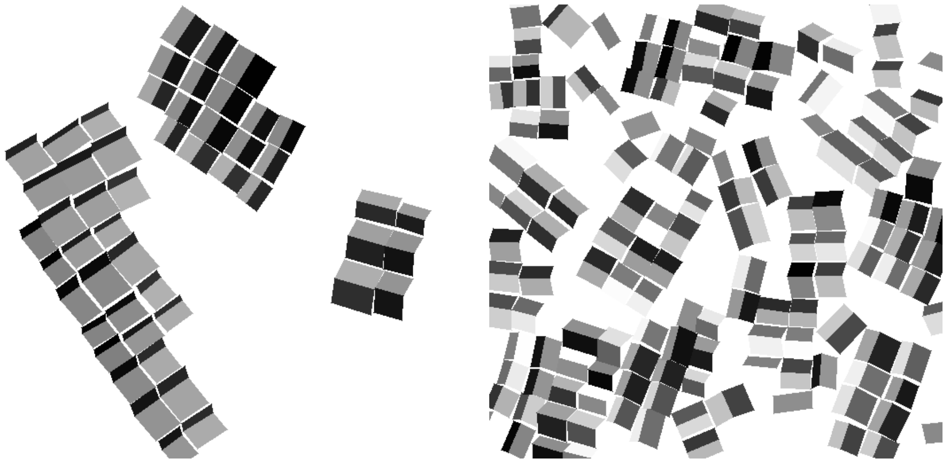


Figure 16: Two configurations obtained with $\beta = 0.0001$ (left handside) and $\beta = 0.01$ (right handside)

8.4 Important remark about the computation of the acceptance ratio

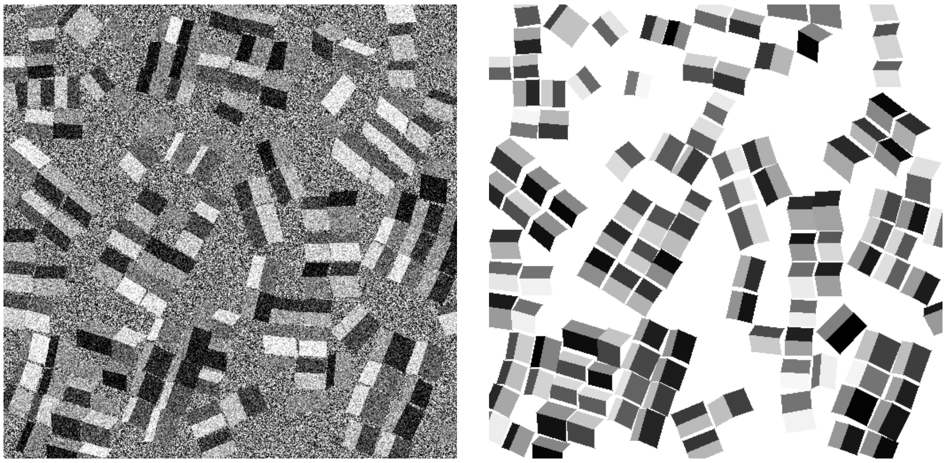
It is essential to notice that the computation of the normalisation constant of the density h is not necessary to achieve the computation of the acceptance ratio. Indeed, we can verify the expression of the acceptance ratio only involves ratios such as $\frac{h(y)}{h(x)}$.

Moreover, according to (9), the acceptance ratio only depends on the added, removed or modified object and on the objects in relation with it and not on the whole configuration. Thus the computation of the ratio $\frac{h(y)}{h(x)}$ is simple and quite fast.

9 Results

9.1 Simulation of the prior model

First we only simulate the prior model without the data model (see Fig. 16).



Left : synthetic image Right : result of the detection

Figure 17: Detecting building in synthetic images

9.2 Data fitting term

9.2.1 Synthetic images

Then, we work on synthetic data (see Fig. 17). This image has been obtained by a simulation of the prior model. We add a Gaussian noise to the rooves and the ground is represented by a uniform noise on the gray levels.

9.2.2 Real images

The aerial photos were provided by the "Institut Géographique National" (French Mapping Institute) and were obtained by a digital camera at a resolution of 0.2m. The viewpoint parameters are given with each image. The result of Figure 18 has been obtained with a stereoscopic couple (i.e. only two images).

10 Prospects

10.1 Prior model

Obviously the geometrical model does not fit any kind of building and it must be refined . For instance, one could define several types of buildings with different

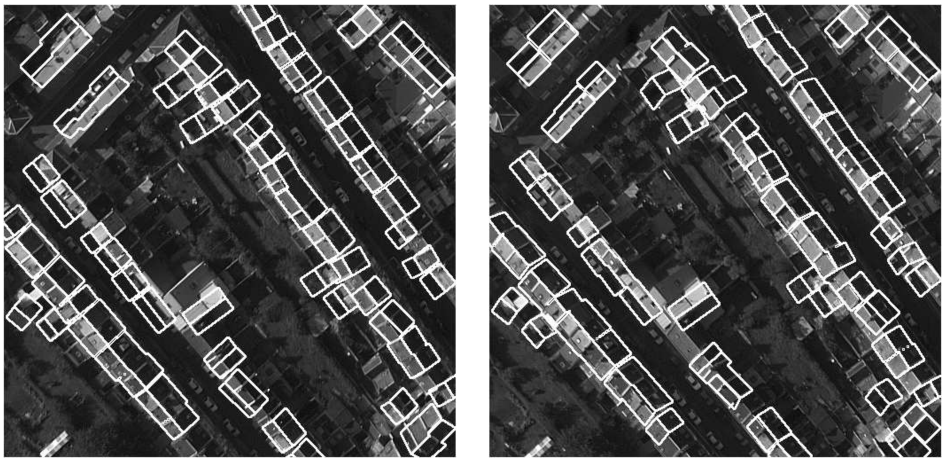


Figure 18: Real images : obtained configuration projected on the left and the right images

geometrical models and add a move which would allow to switch between the different models (an example of this kind is shown in [14]).

Furthermore, the prior model involves only pairwise interaction but one could also use singlewise potential to favour given shapes or, on the contrary, bigger cliques to manage long alignments.

10.2 Data model

There remain a lot of things to do. For instance :

- one can take the occlusions into account,
- one can use a light model.

10.3 Simulation

The proposed algorithm is rather slow. Therefore it is necessary to find some ways to accelerate the convergence.

The first idea could be to propose more relevant moves. Indeed, one could guide the propositions thanks to the image. For instance, it is more efficient to propose buildings which are already located on edges. As we have proposed moves of birth/death in a neighbourhood of a building to fit the prior model, we should also propose moves which fit the data.

We could also use the simulation to estimate some parameters. Indeed, in this work, all the parameters (especially those defining the potentials) were found by trial and error. The estimation could allow us to find the optimal values of the parameters more efficiently. One can refer to [4],[19] for more information on the subject of parameter estimation.

Another possible improvement could be the use of more sophisticated techniques such as the one presented in [20], [21], where the rejection of the propositions can be delayed.

References

[1] D. Stoyan, W. S. Kendall, J. Mecke, "Stochastic Geometry and its Applications", John Willey and Sons, 1987.

[2] D.S. Carter, P.M. Prenter, "Exponential spaces and counting processes",*Z. Wahr. verw. Geb.* 21,1-19,1972.

[3] C. J. Geyer, J. Moller, "Simulation and likelihood inference for spatial point process", *Scandinavian Journal of Statistics, Series B*, 21, 359-373, 1994.

[4] C. J. Geyer, "Likelihood Inference for Spatial Point Processes". In O. E. Barndoff-Nielsen, W. S. Kendall and M. N. M van Lieshout, editors, *Stochastic Geometry, Likelihood and Computation*, Chapmann and Hall, London, 1998.

[5] P. Green, "Reversible Jump MCMC Computation and Bayesian Model Determination", *Biometrika* 82, 711-732, 1995.

[6] P. Green,"MCMC in image analysis". In W. R. Gilks, S. Richardson and D. J. Spiegelhalter, editors, *Markov Chain Monte Carlo in Practice*, pp. 381-399, Chapman and Hall, London, 1996.

[7] L. Tierney, "Markov Chains for Exploring Posterior Distributions", *Tech. Report. No. 560* University of Minnesota, March 15, 1994.

[8] C. Robert and G. Casella, "Monte Carlo Statistical Methods", Springer-Verlag, New York, 1999.

[9] C. Robert, "Méthodes de Monte Carlo par Chaînes de Markov", Ed. Economica, 1996.

[10] B. D. Ripley, "Modelling Spatial Patterns", *Journal of the Royal Statistical Institute, Series B* 39, pp. 172-212, 1997.

[11] B. D. Ripley, F. P. Kelly, "Markov Point Processes", *J. London Math. Soc.*, 15, 188-192, 1977

[12] A. J. Baddeley, M. N. M. van Lieshout, "Recognition of overlapping objects using Markov spatial processes", *CWI Report BS-R9109*, 1991.

[13] L. Tierney (1994), "A note on Metropolis Hastings kernels for general state spaces", *Tech. Rep. 606, School of Statistics, U. of Minnesota.*, 1995.

[14] H. Rue, M.A. Hurn, "Bayesian object identification", *Biometrika* 86, pp. 649-660, 1999.

[15] R. Stoica, X. Descombes, J. Zerubia, "A Markov Point Process for Road Extraction in Remote Sensed Images", *INRIA Research report* No 3923, 2000.

[16] M. Imberty, X. Descombes, "Simulation de processus objets : Etude de faisabilité pour une application à la segmentation d'image", *INRIA Research report* No 3881, 2000.

[17] R. Waagepetersen, "Contributions to the Statistical Modelling of Image Data and Spatial Point Patterns", *PhD Thesis, DMS, University of Aarhus*, 1997.

[18] O. Faugeras, S. Laveau, L. Robert, G. Csurka, C. Zeller, "3-D Reconstruction of Urban Scenes from Sequences of Images", in Automatic Extraction of Man-Made Objects from Aerial and Space Images, Edited by A. Gruen, O. Kuebler and P. Agouris, Birkhäuser, 1995

[19] C. J. Geyer, E. A. Thompson, "Constrained Monte Carlo Maximum Likelihood for Dependent Data", *Journal of the Royal Statistical Institute, Series B* 54, No. 3, pp.657-699,1992.

[20] P. J. Green, A. Mira, "Delayed Rejection in Reversible Jump Metropolis-Hastings", *Mathematics Research Report* n. S-01-99, University of Bristol, 1999.

[21] L. Tierney, A. Mira, "Some adaptive Monte Carlo methods for Bayesian inference", *Statistics in Medicine* 18, pp. 2507-2515, 1999.

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

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

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