Texture Representation by Geometric Objects using a Jump-Diffusion Process

Florent Lafarge, Georgy Gimel'Farb

Texture Representation by Geometric Objects using a Jump-Diffusion Process

Florent Lafarge and Georgy Gimel'farb
Dept. of Computer Science, University of Auckland
Private Bag 92012, Auckland 1142, New Zealand
{flaf006,ggim001}@cs.auckland.ac.nz

Abstract

Our goal is to represent images in terms of geometric objects acting as primitive elements of an image description. Similar representations obtained by stochastic marked point processes have already led to convincing image analysis results but suffer from serious drawbacks such as complex and unstable parameter tuning, large computing time, and lack of generality. We propose an alternative descriptive model based on a Jump-Diffusion process which can be performed in shorter computing times and applied to a variety of applications without changing the model or modifying the tuning parameters. In our approach, a probabilistic Gibbs model is adapted to a library of geometric objects and is sampled by a Jump-Diffusion process in order to closely match an underlying texture. Experiments with natural textures and remotely sensed images show good potentialities of the proposed approach[1].

1 Introduction

Shape extraction is a well known computer vision problem which has been addressed by various approaches. Deformable models such as parametric or level sets based active contours are particularly efficient to extract or track curved shapes and have been successfully used, for example, for recognition of organs in medical imaging. Alternative approaches that use stochastic models and random sampling of geometric objects are better adapted to extraction of rectilinear shapes. Stochastic models often involve jump samplers which allow to deal with state spaces of variable dimension. The most known jump sampler is the Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm [6]. It is particularly efficient for recognition of objects with variable numbers of parameters in large configuration spaces, such as 3D reconstruction [2, 10] or texture modelling [19].

Marked point process based models are among the most efficient stochastic approaches and have already obtained convincing results in various imaging applications such as building extraction [13], road network detection [9], or tree crown extraction [14]. The marked point processes, detailed in [16], exploit random variables whose realisations are configurations of geometrical objects (for example, rectangles [13], line segments [9], or ellipses [14]). An energy is associated with each object configuration, and the global minimum of this energy is searched for by using the conventional simulated annealing [12]

[1]The first author is grateful to the French Defence Agency (DGA) for financial support. Both the authors thank the French Mapping Agency (IGN) and the French Forest Inventory (IFN) for the aerial images.

coupled with the birth-and-death sampler [4]. Such processes allow to describe complex spatial interactions between the objects. Image representations provided by the marked point process based models are particularly good for complex textured images. However, these models suffer from the following three major drawbacks:

Lack of generality Each model is associated with a specific application, and a marked point process is limited to a single type of objects having simple geometric shape. Moreover, complexity of interactions between the objects defined in the model makes impossible to generalise each particular model to another application.

Trial-and-error parameter tuning Many parameters (up to ten in most of cases) are to be used to define the interactions. They are tuned by trial and errors since parameter estimation techniques which efficiently work with such large configuration spaces simply do not exist.

Too long computational time Although proposition kernels are developed to speed up the process, the birth-and-death sampler remains very slow, especially at low temperature.

This paper presents an alternative to the marked point processes. Our aim is not to provide results as accurate as those obtained by models based on the marked point process, but to propose a more general stochastic process which can produce target results in shorter time and can be applied to a large range of applications without modifying model and tuning parameters. The proposal involves the following two modifications of the marked point process:

Sampling several types of objects and limiting their interactions: In order to extend the level of generality, the process must jointly sample different types of geometric objects (*e.g.* linear objects such as segments and areal ones such as polygons and circles). The interactions between these objects must also be reduced and simplified in order to strongly decrease the number of tuning parameters.

Introducing a diffusion dynamic: The diffusion dynamic would allow us to significantly speed up the convergence of the process. The marked point process based models cannot use such a dynamic due to complexity of their energy functions (usually gradients of these functions do not satisfy the Lipschitz continuity condition). The Jump-Diffusion processes introduced by Grenander *et al.*[7] represent a class of random samplers which efficiently mix both the jump and diffusion dynamics.

The paper is organised as follows. Section 2 presents a Gibbs energy model adapted to different types of geometric objects. The model is sampled by a Jump-Diffusion process detailed in Section 3. Experimental results for texture description problems are given in Section 4. Basic conclusions are outlined in Section 5.

2 Image representation model

By way of illustration, we restrict ourselves to a simple object set (or library) of seven geometric patterns (see Fig. 1). Segments, lines, and line ends are specific to linear structures whereas rectangles, bands, band ends, and circles correspond to areal descriptors.

All the objects have between three and five control parameters, the positional coordinates (x_c, y_c) of the object's centre being common to all the objects. Other parameters depend on the object types (e.g. radius for circles; length, width, and orientation for bands and rectangles; or length and orientation for lines and segments). The parameters are defined in both continuous and discrete domains. This set includes basic objects used in the known marked point process based models and thus is sufficient to produce detailed representations of a large range of scenes in terms of their linear and areal components.

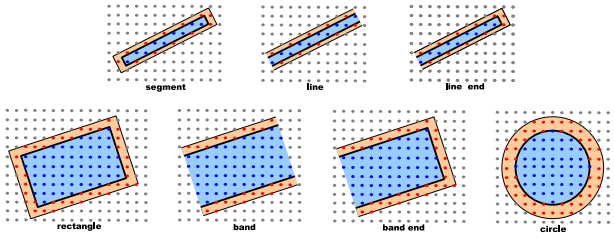

Figure 1: The chosen library of linear and areal geometric objects.

2.1 Gibbs energy

The number of objects in any particular scene is unknown, and the objects have different numbers of parameters. Thus, the configuration space $\mathscr{C}$ of our problem is defined as an union of subspaces $\mathscr{C}_k$, each subspace containing fixed numbers of objects of each type. A probability distribution μ on the configuration space $\mathscr{C}$ is defined as a combination of μ_k distributions on the subspaces $\mathscr{C}_k$. We assume unnormalized distributions μ_k on $\mathscr{C}_k$ have Gibbs densities of the form $e^{-E_k(x)}$ where E_k is a Gibbs energy associated with the configuration subspace $\mathscr{C}_k$.

The energy E_k takes into account both the coherence $D_k(x)$ between the objects and the image data and the adjacency constraints $R_k(x)$ for positioning of the objects with no overlaps:

$$E_k(x) = D_k(x) + R_k(x); \quad x \in \mathscr{C}_k \tag{1}$$

2.1.1 The data coherence term

$D_k(x)$ accumulates the local energy associated with each object x_i of the configuration x:

$$D_k(x) = \sum_i d(x_i) \tag{2}$$

where $d(x_i)$ is a measure of coherence of the object x_i with respect to the data (*i.e.* an image). This measure $d(.)$ must satisfy two important conditions:

- It must be **independent of the object type**. In particular, the object area must be taken into account in order to not favour linear or areal object types.
- It must allow to **select "attractive" objects**, *i.e.* the well-fitted objects having a negative local energy. This feature is very important in the models using birth-and-death processes [16, 13, 9] since it partly defines the object density in the scene.

The function we propose is derived from the Mahalanobis distance and includes a threshold θ_{attr} that makes some objects attractive if the function is negative:

$$d(x_i) = \begin{cases} \sqrt{\frac{\sigma_{\text{in}}^2 + \sigma_{\text{out}}^2 + \varepsilon}{S(m_{\text{in}} - m_{\text{out}})^2}} - \theta_{\text{attr}} & \text{if } m_{\text{in}} \neq m_{\text{out}} \\ \infty & \text{otherwise} \end{cases} \tag{3}$$

Here, m_{in} and m_{out} represent the mean of pixel intensities inside and outside the object respectively[2] (*i.e.* the blue and red areas on Fig. 1), σ_{in} and σ_{out} denote the associated standard deviations, S is the whole inside and outside area, and $\varepsilon > 0$ is an infinitesimal value allowing $d(.)$ to be derivable. The threshold θ_{attr} allows to select the attractive objects and tune the sensitiveness of the data fitting. This measure of coherence is based on homogeneity criteria inside and outside the object. The computation time of this measure is very short, but this function is not optimal for noisy data. Nonetheless, it produces better experimental results than other measures, *e.g.* based on the Bhattacharya distance.

2.1.2 The adjacency constraint

$R_k(x)$ follows from the unique prior knowledge and is necessary for developing a general model of non-overlapping objects. Other types of interactions such as inter-connections or mutual alignments of the objects could be also introduced. However, our aim is to minimise the number of tuning parameters in the model since it is a critical problem for the marked point processes. This term is expressed as follows:

$$R_k(x) = \sum_{x_i, x_j \in x} (e^{\kappa g(x_i, x_j)} - 1) \tag{4}$$

where $g(x_i, x_j)$ taking values in $[0, 1]$ quantifies the mutual overlap between the objects x_i and x_j, and κ is a big positive real value ($\kappa >> 1$) which strongly penalises the overlaps (in our experiments, $\kappa = 100$). Under small overlaps between two objects, this prior will weakly penalise the global energy. But if the overlapping is high, this prior will act as an hardcore (*i.e.* the prior energy takes a very high value), and the configuration will be practically banned.

3 Jump-Diffusion sampler

The search for an optimal configuration of objects is performed using the Jump-Diffusion process introduced by Grenander *et al.* [7]. It has been used in various applications such as target tracking [15] and image segmentation [8]. This process combines the conventional Markov Chain Monte Carlo (MCMC) algorithms [12, 6] and the Langevin equations [3]. Both dynamics play different roles in the Jump-Diffusion process: the former performs reversible jumps between the different subspaces $\mathscr{C}_k$, whereas the latter conducts stochastic diffusion within each continuous subspace. The global process is controlled by a relaxation temperature T depending on time t and approaching zero as t tends to infinity. Simulated annealing theoretically ensures convergence to the global optimum from any

[2] By taking $m_{\text{in}} > m_{\text{out}}$ or $m_{\text{in}} < m_{\text{out}}$ instead of $m_{\text{in}} \neq m_{\text{out}}$ in the definition domain of $d(.)$, we can modify the measure in order to introduce radiometric information and favor respectively bright or dark objects with respect to the background. This variant of $d(.)$ will be used for tree crown and building extraction experiments.

initial configuration using a logarithmic decrease of the temperature. In practice, we use a faster geometric decrease which gives an approximate solution close to the optimum. The simulated annealing parameters such as the initial temperature are estimated using the approach of White [18]. The diffusions are interrupted by jumps following a discrete time step Δt (in practice, $\Delta t = 50$). At the very low temperature, the diffusion process plays a more important role: the time step is increased ($\Delta t = 100$) to speed up the convergence.

3.1 Jump dynamic

Reversible jumps between the different subspaces are performed according to families of moves called proposition kernels and denoted by Q_m. The jump process performs a move from an object configuration $x \in \mathscr{C}_k$ to $y \in \mathscr{C}_{k'}$ according to a probability $Q_m(x \to y)$. Then, the move is accepted with the following probability:

$$\min\left(1, \frac{Q_m(y \to x)}{Q_m(x \to y)} e^{-\frac{(E_{k'}(y) - E_k(x))}{T}}\right) \tag{5}$$

We use two different families of moves in order to jump between the subspaces.

Birth-and-death kernel Q_1***:*** This kernel allows for adding or removing an object from a current object configuration. These transformations corresponding to jumps into the spaces of higher (birth) and lower (death) dimension are theoretically sufficient to visit the whole configuration space. In practice, we choose to add or remove an object following a Poisson distribution. If an object is added, its type is randomly chosen and its parameters are chosen according to uniform distributions over the parameter domains. The computation of this kernel is detailed in [16, 4].

Switching kernel Q_2***:*** This kernel allows to switch the type of an object (*e.g.* a circle by a rectangle). Contrary to the previous kernel, this move do not change the number of objects in the configuration. However, the number of parameters can be different (*e.g.* three parameters for a circle substituted by five parameters for a rectangle). This kernel is based on the creation of bijections between the different types of objects. The computation of this kernel is detailed in [6].

Usually the jump processes [13, 9, 14, 10] use a perturbation kernel that allows them to explore each subspace by modifying only parameters of the objects. In our case, this kernel is substituted by a diffusion dynamic which is clearly faster since the exploration of the subspace is directed by the energy gradient.

3.2 Diffusion dynamic

The diffusion process between jumps controls the dynamics of the object configuration in their respective subspaces. Stochastic diffusion (or Langevin) equations driven by Brownian motions $dB(t)$ with temperature T are used to explore the subspaces $\mathscr{C}_k$. If $x(t)$ denotes the variables at time t, then

$$dx(t) = -\frac{dE_k(x)}{dx}dt + \sqrt{2T(t)}dw_t \tag{6}$$

where $dw_t \sim N(0, dt^2)$. At high temperature ($T >> 0$), the Brownian motion is useful in avoiding trapping in local optima. At low temperature ($T << 1$), the role of the Brownian motion becomes negligible and the diffusion dynamic acts as a gradient descent.

4 Experimental results

4.1 Texture representation

The proposed method has been tested on a number of selected natural textures in order to evaluate its potentialities of representing various types of images by geometric objects. The obtained results (some of them are presented in Fig. 2) are quite promising. Various spatially homogeneous and heterogeneous textures are successfully represented even with a chosen simple set of objects. Some textures having spatially variant illumination and reflectance (see *e.g.* the metal grid and tile roof examples in Fig. 2) are usually difficult to describe, and often require specific advanced techniques such as [1]. Our method is particularly interesting for representing such textures since the fitting of objects do not depend of illumination effects.

Fig. 3 presents both the result obtained from an image containing five different textures and the evolution of the object configuration during the jump-diffusion process. This result showing five various object layouts underlines interesting potentialities for texture discremination. At the beginning of the algorithm, *i.e.* when the temperature is high *(red)*, the process explores the subspaces and favors configurations with a low energy. At this exploration stage, the jump dynamic plays an important role by specifying both the number and the type of objects. At low temperature *(blue)*, the object configuration belongs to a subspace being close to the optimal one, and the number of objects in the scene does not evolve very much. The diffusion dynamic is mainly useful at this stage in order to perform a detailed adjustment of the object parameters. This dynamic is clearly faster than a single jump process with a perturbation kernel since the exploration is directed by the gradient of the energy (and not by a random search). Graphs in Fig. 3 describe how the energy and the number of objects change in function of the number of iterations.

Even if the objects are disconnected (see *e.g.* brick wall or hair in Fig. 2), the representation is detailed enough to be useful in solving texture description and recognition problems. In particular, it would be interesting to combine such object-based representations with Gibbs Markov random field models which are mostly used on the pixel-wise intensities [5] and thus cannot explicitly take into account shapes and relative locations of depicted characteristic objects. In order to deal with more complicated textures and have a description level similar to filter bank methods such as [17], the object library has to be extended and new relevant shapes, especially curved shapes, should be introduced. Moreover, it will be necessary to develop more general energy functions taking into account, in particular, typical object deviations and noise in the textures (see the stone ornament and rose results in Fig. 2 which are quite limited in term of description).

4.2 Remote sensing applications

A similar approach detailed in [11] has been tested on different remote sensing problems such as bird detection, tree crown extraction, road network detection, and building extraction, with the data used in the marked point process based methods [13, 9, 14]. Although the obtained results are generally less accurate than those obtained by the specialised marked point process based methods, the proposed general process allows to deal with various remote sensing problems in much shorter time and without modifying the model and tuning parameters. Some results are presented on Fig. 4. For the tree crown extraction, the main goal is to count trees on large forest scenes. Although the shapes of trees

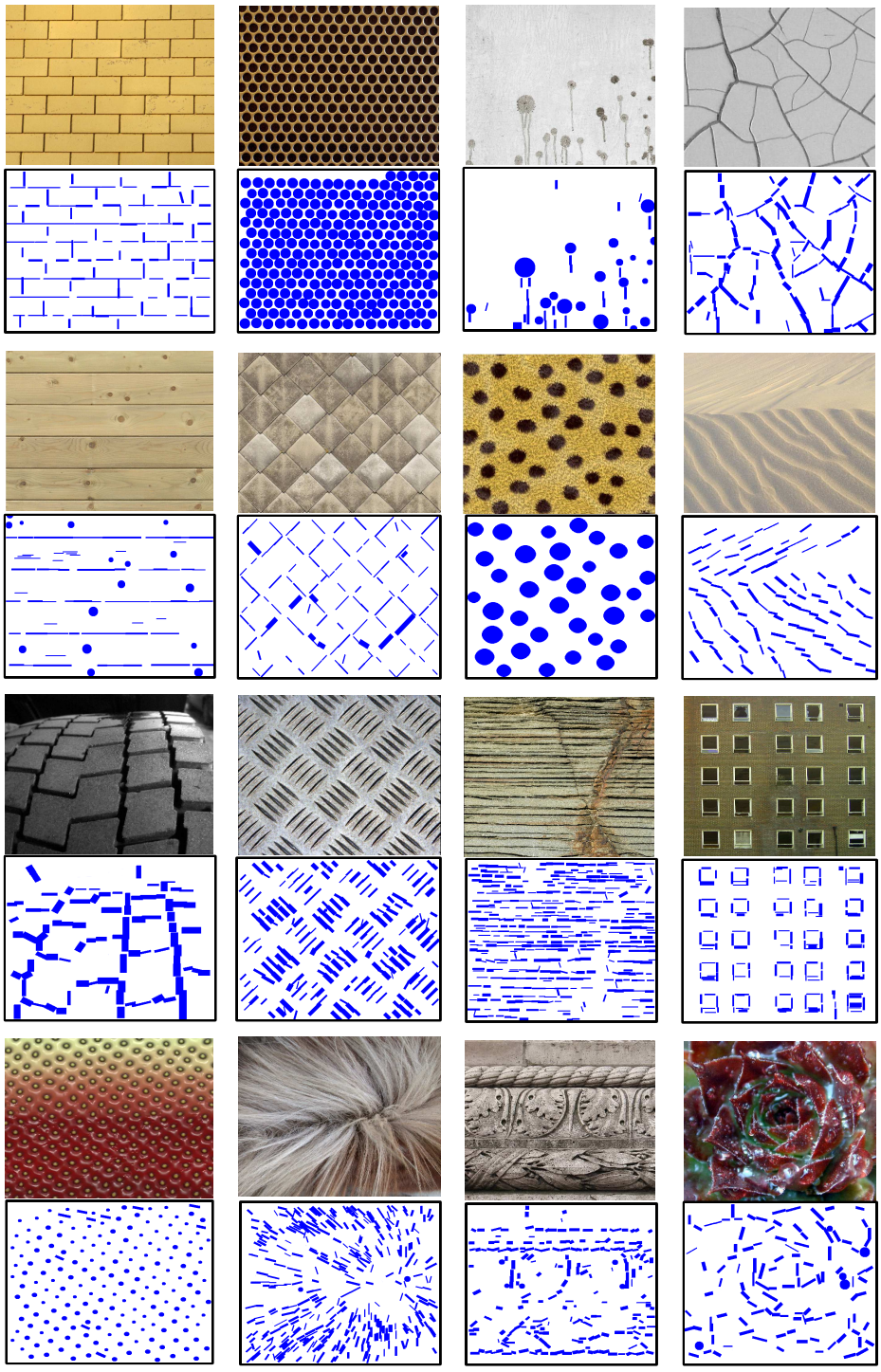

Figure 2: Examples of our representation of textures from `http://www.cgtextures.com` in terms of geometrical objects.

are roughly approximated by circles and rectangles, all the trees are accurately detected. The accuracy of the tree locations is practically the same as obtained by a marked point process [14] with elliptical objects to represent the trees.

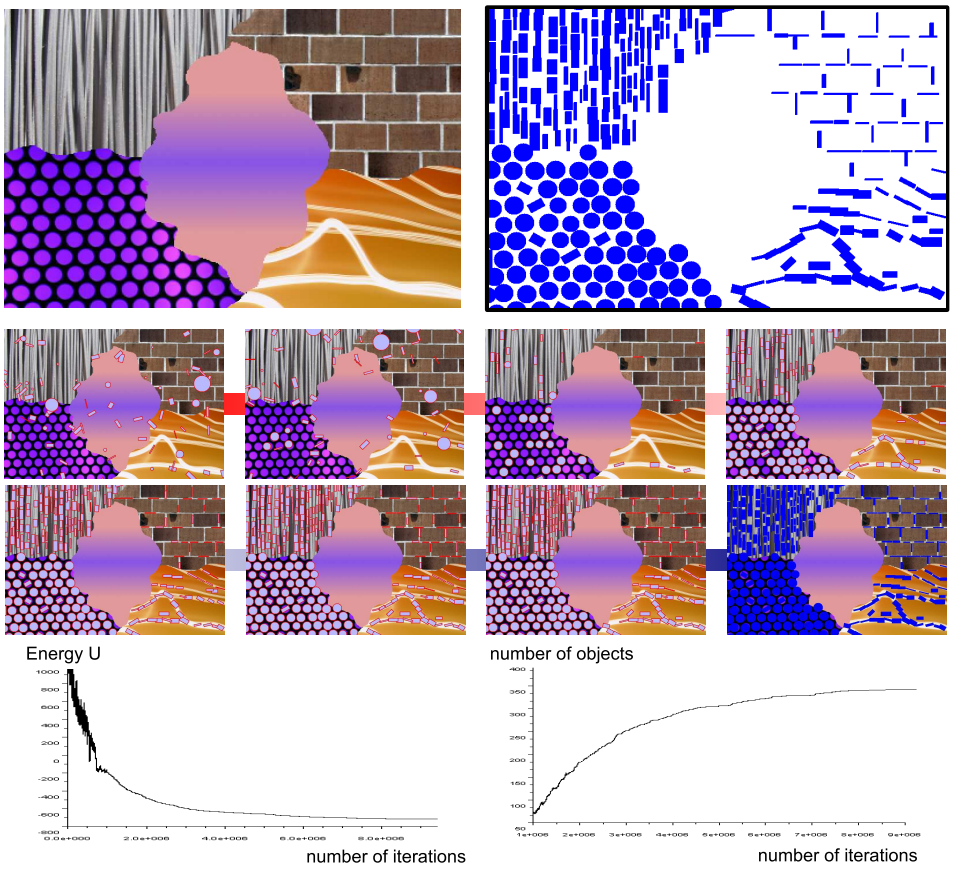

Figure 3: Texture representation: original image and result *(top)*; evolution of the object configuration during the jump-diffusion process - configurations from the initial temperature (red) to the final one (blue) *(middle)*; energy and number of objects graphs in function of the number of iterations during the critical phase *(bottom)*.

The road network and building extraction results cannot be considered as a final representation since the detected objects are not connected (contrary to [9] or [13] where complex interactions had been defined to link the objects). However, the objects found are mainly lines and bands which are globally well fitted to roads and buildings and provide a rough pattern of the target structures. These object layouts are sufficiently informative to make it possible to extract the global network on the basis of their subsequent analysis. For example, one could use post-processing based on a vectorisation principle to connect the objects found. The building extraction is convincing but the object localisation remains very rough compared to the one obtained in [13]. However, our method is clearly

faster: 30 minutes *vs* 2 hours on a 0.3 km^2 dense urban area using a 3 GHz processor.

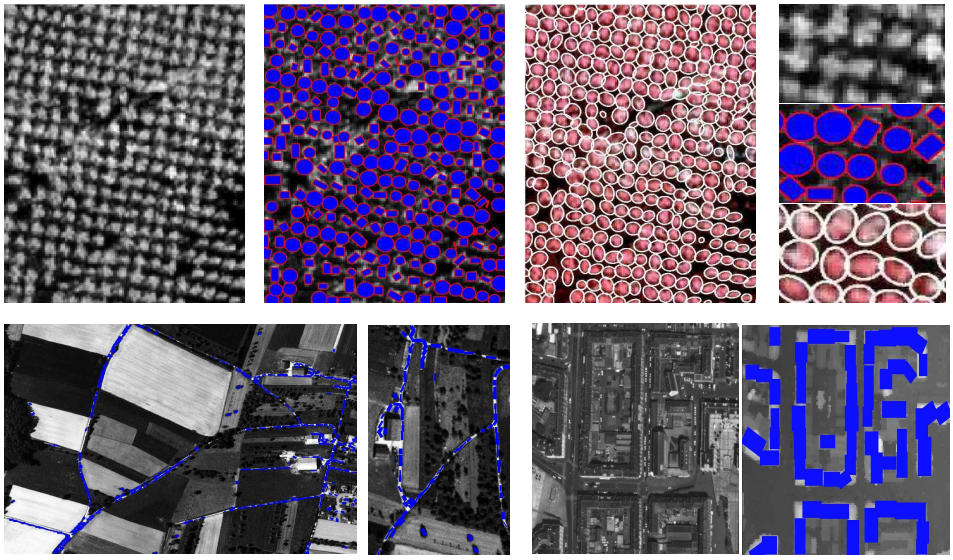

Figure 4: Remote sensing applications from aerial images *(from top to bottom, and left to right)*: Tree crown extraction with the original image ©IFN, our result, the result obtained in [14], and crops *(top)*; a road network extraction result and its cropped part, aerial image of a urban scene ©IGN and the buildings extracted from the associated DEM *(bottom)*.

5 Conclusions

We have proposed a new approach for the representation of images in terms of simple, or primitive geometric objects. The approach possesses several important characteristics comparing to its conventional counterparts based on the marked point processes and models. It is more general and works efficiently on various applications without modifying the model and tuning parameters. Moreover, the optimisation technique based on the jump-diffusion process allows to obtain shorter computational time compared to the classical jump processes. However, the proposed process is limited by the content of the object library (the current set in Fig. 1 cannot in principle provide relevant representations of complex textures including curved object shapes and noise). In the future, it could be particularly interesting, first, to extend the object library and, secondly, to develop models and techniques for automatic selection of relevant objects from a given collection of training image before the use of the above jump-diffusion process.

References

[1] M.J. Chantler, M. Petrou, A. Penirsche, M. Schmidt, and G. McGunnigle. Classifying surface texture while simultaneously estimating illumination direction. *IJCV*,

62(1-2):83–96, 2005.

[2] A.R. Dick, P.H.S. Torr, and R. Cipolla. Modelling and interpretation of architecture from several images. *IJCV*, 60(2):111–134, 2004.

[3] S. Geman and C.R. Huang. Diffusion for global optimization. *SIAM Journal on Control and Optimization*, 24(5):131–143, 1986.

[4] C.J. Geyer and J. Moller. Simulation and likelihood inference for spatial point processes. *Scandinavian Journal of Statistics*, Series B(21):359–373, 1994.

[5] G. Gimel'Farb. *Image Textures and Gibbs Random Fields*. Springer, 1999.

[6] P.J. Green. Reversible Jump Markov Chains Monte Carlo computation and Bayesian model determination. *Biometrika*, 57:711–732, 1995.

[7] U. Grenander and M.I. Miller. Representations of Knowledge in Complex Systems. *Journal of the Royal Statistical Society*, 56(4):1–33, 1994.

[8] F. Han, Z.W. Tu, and S.C. Zhu. Range Image Segmentation by an Effective Jump-Diffusion Method. *IEEE PAMI*, 26(9):1138–1153, 2004.

[9] C. Lacoste, X. Descombes, and J. Zerubia. Point processes for unsupervised line network extraction in remote sensing. *IEEE PAMI*, 27(10):1568–1579, 2005.

[10] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny. Building reconstruction from a single dem. In *IEEE CVPR*, Anchorage, United States, 2008.

[11] F. Lafarge, G. Gimel'farb, and X. Descombes. A geometric primitive extraction process for remote sensing problems. In *ACIVS*, Juan-Les-Pins, France, 2008.

[12] M. Metropolis, A.W. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. of Chemical Physics*, 21:1087–1092, 1953.

[13] M. Ortner, X. Descombes, and J. Zerubia. Building outline extraction from Digital Elevation Models using marked point processes. *IJCV*, 72(2):107–132, 2007.

[14] G. Perrin, X. Descombes, and J. Zerubia. Adaptive simulated annealing for energy minimization problem in a marked point process application. In *EMMCVPR*, St Augustine, United States, 2005.

[15] A. Srivastava, M. Miller, and U. Grenander. Multiple target direction of arrival tracking. *IEEE SP*, 43(5):282–285, 1995.

[16] M.N.M Van Lieshout. Markov point processes and their applications. In *Imperial College Press*. London, 2000.

[17] M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *IJCV*, 62(1-2):61–81, 2005.

[18] S.R. White. Concepts of scale in simulated annealing. In *Proc. IEEE ICCD*, 1984.

[19] Y.N. Wu, S.C. Zhu, and C. Guo. Statistical modeling of texture sketch. In *ECCV*, Copenhagen, Denmark, 2002.