

INTERNATIONAL JOURNAL OF COMPUTER VISION AND IMAGE PROCESSING

April-June 2011, Vol. 1, No. 2

Table of Contents

RESEARCH ARTICLES

- 1 A Marked Point Process Model Including Strong Prior Shape Information Applied to Multiple Object Extraction From Images**
Maria Kulikova, INRIA Sophia-Antipolis, France
Ian Jermyn, INRIA Sophia-Antipolis, France
Xavier Descombes, INRIA Sophia-Antipolis, France
Elena Zhizhina, Institute of Information Transmission Problems (IITP), Russia
Josiane Zerubia, INRIA Sophia-Antipolis, France

- 13 A Robust Embedding Scheme and an Efficient Evaluation Protocol for 3D Meshes Watermarking**
Saoussen Ben Jabra, RIADI Laboratory, Team of Research SIIVA, High Institute of Computer Science, University of Tunis El Manar, TUNISIA
Ezzeddine Zagrouba, RIADI Laboratory, Team of Research SIIVA, High Institute of Computer Science, University of Tunis El Manar, TUNISIA

- 30 Direct 3D Information Determination in an Uncalibrated Stereovision System by Using Evolutionary Algorithms**
Alain Koch, Université de Bourgogne, France
Albert Dipanda, Université de Bourgogne, France
Claire Bourgeois-République, Université de Bourgogne, France

- 42 Construction of 3D Triangles on Dupin Cyclides**
Bertrand Belbis, Université de Bourgogne, France
Lionel Garnier, Université de Bourgogne, France
Sebti Foufou, Université de Bourgogne, France, and Qatar University, Qatar

- 58 Evaluation Approach of Arabic Character Recognition**
Hanan Aljuaid, University Technology Malaysia, Malaysia
Dzulkifli Mohamad, University Technology Malaysia, Malaysia
Muhammad Sarfraz, Kuwait University, Kuwait

A Marked Point Process Model Including Strong Prior Shape Information Applied to Multiple Object Extraction From Images

Maria Kulikova, INRIA Sophia-Antipolis, France

Ian Jermyn, INRIA Sophia-Antipolis, France

Xavier Descombes, INRIA Sophia-Antipolis, France

Elena Zhizhina, Institute of Information Transmission Problems (IITP), Russia

Josiane Zerubia, INRIA Sophia-Antipolis, France

ABSTRACT

Object extraction from images is one of the most important tasks in remote sensing image analysis. For accurate extraction from very high resolution (VHR) images, object geometry needs to be taken into account. A method for incorporating strong yet flexible prior shape information into a marked point process model for the extraction of multiple objects of complex shape is presented. To control the computational complexity, the objects considered are defined using the image data and the prior shape information. To estimate the optimal configuration of objects, the process is sampled using a Markov chain based on a stochastic birth-and-death process on the space of multiple objects. The authors present several experimental results on the extraction of tree crowns from VHR aerial images.

Keywords: Active Contour, Marked Point Process, Multiple Birth-and-Death Dynamics, Multiple Object Extraction, Shape Prior

INTRODUCTION

Object detection from optical satellite and aerial images is one of the most important tasks in remote sensing image analysis. The problem arises in many applications, both civilian and military,

e.g. tree counting and species classification for biomass or biodiversity estimation; and bird counting for monitoring population changes. Nowadays the resolution of aerial images is approaching a few centimeters. At this level of resolution, the geometry of objects is clearly visible, and needs to be taken into account for accurate object extraction.

DOI: 10.4018/ijcvip.2011040101

Stochastic point process models are known for their ability to include this type of information. A probability distribution is defined on the space of configurations composed of multiple objects which depends on the relation between the objects and data, and on the configuration of individual objects as well as their joint relations. The extracted objects are then those in the ‘optimal’ configuration, which is usually estimated using maximum a posteriori (MAP) estimation.

In previous work, Marked Point Process (MPP) models have been used for the extraction of road networks (Lacoste, Descombes, & Zerubia, 2005), buildings (Ortner, Descombes, & Zerubia, 2007), trees (Perrin, Descombes, & Zerubia, 2005) and birds (Descamps, Descombes, Béchet, & Zerubia, 2009), from images of more than 50 cm/pixel resolution. At that level the objects have a simplified geometrical shape and were thus represented using simple shape objects, *e.g.* discs, ellipses, or rectangles.

Recently we lifted this restriction without increasing the dimension of the space of a single object (Kulikova, Jermyn, Descombes, Zerubia, & Zhizhina, 2010). A single object was represented by its boundary, a closed curve, but the set of possible single objects (*i.e.* boundaries) was defined not *a priori*, but by the image data and a single-object version of the model. A probability distribution was then defined on the configuration space of an unknown number of objects.

This approach is well suited to scenes composed of objects that do not vary too much in shape and size within a class, and that have smooth enough boundaries. In this case, accuracy in the number of extracted objects can be achieved by favouring the smoothness term that controls, as well, the length of the curve. The model can thus to some degree separate objects that overlap, but this leads to imprecision in delineating objects.

The aim of this paper is to incorporate into the single-object model prior knowledge about the shape of the objects sought, in order to deal with overlapping objects with complex shapes

without losing their geometric details and without significantly increasing the computational complexity of estimation.

Our work can also be viewed as an extension of the active contour methodology (Kass, Witkin, & Terzopolous, 1998) to cases in which the number of objects is unknown *a priori*, and where shape prior information is incorporated as well. Much work has been already done based on the active contour approach. Some of this work, *e.g.* (Caselles, Kimmel, & Sapiro, 1997), includes only weak shape information, essentially smoothness, but can in principle detect multiple (although not overlapping) objects using the level set representation (Osher & Fedkiw, 2003; Sethian, 1999). Other work includes much stronger prior information about shape (Cremers, Tischauser, Weickert, & Schoerr, 2002; Cremers, Osher, & Soatto, 2006b; Joshi, & Srivastava, 2009; Leventon, Grimson, & Fougereas, 2000), but the method used to do this means that it is difficult to treat an unknown number of objects (and in practice only single objects are treated). Cremers et al. (2006a) treat the problem of image segmentation into connected components each of which corresponds to one class of a number of distinct classes of objects, but only one object can be found in each connected component.

Additionally, the results obtained by these methods may be very dependent on the initial configuration, since the algorithm typically used is deterministic gradient descent. Storvik (1994) and Juan et al. (2006) respectively use a Markov Chain Monte Carlo (MCMC) approach and stochastic partial differential equation (SPDE) optimization techniques to minimize active contour energies, but the algorithms make only local changes to the contour at each iteration, and thus preserve object number. Srivastava and Jermyn (2009) use a stochastic algorithm together with an energy including strong prior shape information for the extraction and classification of objects from 2D point clouds, but only a single object is sought. Tu et al. (2002) use data-driven reversible jump Markov Chain Monte Carlo (RJMCMC) dynamics to solve a problem of general purpose image segmenta-

tion, but our work differs, first, in addressing a specific problem rather than general possibilities; and second, in using birth-and-death rather than RJMCMC dynamics. The advantage of birth-and-death dynamics is their faster convergence, due to the fact that at each of iterations, several objects are simultaneously added to the current configuration without any rejection probability.

The remainder of the article is organized as follows. Firstly, we describe the single-object space and the single-object term in the energy function. Then, we describe the multiple-object space and the full energy, and the sampling and estimation algorithm. Finally, we present experimental results and conclusion.

Single-Objects

As mentioned, the single-object space will not be determined *a priori* to consist of simple geometrical shapes, but rather will be constructed using the image data and a model describing configurations of individual objects. We model individual object boundaries as closed planar curves $\gamma : [0, 2\pi] \rightarrow V \subset \mathbf{R}^2$ lying in the image domain V . The set of closed curves we will consider here consists of ‘star domains’ parameterized by $(x_0, \delta r(t))$, where $x_0 \in \mathbf{R}^2$ and $\delta r : [0, 2\pi] \rightarrow \mathbf{R}$ is a radial variation around a circle $\gamma_c(t)$ of radius $r_0 \in [r_{\min}, r_{\max}]$ centred at $x_0 \in V$, where $t \in [0, 2\pi]$. Then,

$$\gamma(t) = x_0 + \gamma_c(t) + \delta\gamma(t), \quad (1)$$

where $\delta\gamma$ is the radial variation expressed in Euclidean coordinates. More explicitly,

$$\begin{aligned} \gamma(t) &= \\ x_0 + ((r_0 + \delta r(t)) \cos \theta_c(t), (r_0 + \delta r(t)) \sin \theta_c(t)) \\ &= (x_0^x, x_0^y) + (r_0 + \delta r(t)) (\cos(t), \sin(t)). \quad (2) \end{aligned}$$

We suppose that we are given an energy function E defined on a space Γ of these curves (with appropriate restrictions to ensure

that everything is well-defined). This energy function will depend on the image data also, as it will be detailed below in the article.

Given an initial curve $\gamma \in \Gamma$, we can then perform gradient descent to arrive at a local minimum of E , giving a second curve, $\tilde{\gamma} \in \Gamma$. The map $\tau : \Gamma \rightarrow \Gamma$ takes every curve to the local minimum in whose basin of attraction it lies. Now define the space \mathbf{C} to be a set of circles lying in the image domain, with radii $r_0 \in [r_{\min}, r_{\max}]$, parameterized by arc length. The single-object space we consider is $\Gamma_o = \hat{\mathbf{C}}$. The possible single objects are thus locally adapted to the data. Then despite allowing for potentially arbitrarily shaped star domains, the dimension of the single-object space remains small: if we fix the centre of the circle in \mathbf{C} , *i.e.* the ‘point’ in the marked point process, the ‘mark’ is one-dimensional, being equivalent to the circle’s radius.

Single-Object Energy

To define the space Γ_o , we have to define the energy E . In this paper, we define it as a sum of two terms: a term related to the curve, *i.e.* a prior term, and an image term, also called the data term:

$$E(\gamma) = E_{\text{image}}(\gamma) + E_{\text{curve}}(\gamma). \quad (3)$$

The image energy term is defined as a weighted sum:

$$E_{\text{image}}(\gamma) = \lambda_g \int_{[0, 2\pi]} dt n(t) \cdot \nabla I(\gamma(t)) \quad (4)$$

where $n(t)$ is the (unnormalized) outward normal to the curve; I is the image; $G(x) = \frac{(I(x) - \mu)^2}{2\sigma^2}$ and $\bar{G}(x) = \frac{(I(x) - \bar{\mu})^2}{2\bar{\sigma}^2}$; and $R(\gamma)$ is the interior region corresponding to the boundary γ . The first term favours boundaries with high image gradients normal to the boundary. The second term arises from a Gaussian image model with different means

and variances for the interior and exterior of the objects. Both terms are negative where $\gamma(t)$ is well-adapted to the data. The parameters μ , σ , and $\bar{\mu}$, $\bar{\sigma}$, are learned from examples of object and background.

We define $E_{\text{curve}}(\gamma)$ as a sum of two terms:

$$E_{\text{curve}}(\gamma) = E_{\text{smth}}(\gamma) + E_{\text{sh}}(\gamma), \quad (5)$$

We show below that it can be written in the following form:

$$E_{\text{curve}}(\gamma) = \iint_{[0,2\pi]} dt dt' F(t-t') \delta r(t) \delta r(t'). \quad (6)$$

The $E_{\text{smth}}(\gamma)$ term favours boundary smoothness and a uniform parameterisation of the curve:

$$E_{\text{smth}}(\gamma) = \int_{[0,2\pi]} dt |\dot{\gamma}(t)|^2, \quad (7)$$

where $\dot{\gamma}$ is the derivative of γ .

The second term $E_{\text{sh}}(\gamma)$ represents the prior energy associated with the curve shape. It is a quadratic function of δr , which due to invariance to translations of the origin of the curve, is diagonal in the Fourier basis on the circle, and has zero mean except at zero frequency. The latter corresponds to changes in radius, and are absorbed in r_0 . The energy is thus defined by the variance of each Fourier component, $\sigma(k)^2 = \frac{1}{4\pi g(k)}$. The function

$g(k)$ restricts or favours perturbations of the circle with different frequencies k , thereby reflecting the specificities of the shapes of the objects to be detected in the image:

$$E_{\text{sh}}(\gamma) = 2\pi \sum_{k \in \mathbb{Z}} g(k) |\hat{\delta r}(k)|^2, \quad (8)$$

where:

$$\hat{\delta r}(k) = \frac{1}{2\pi} \int_{[0,2\pi]} dt \exp(-ikt) \delta r(t). \quad (9)$$

Figure 1 shows examples of shapes obtained from a circle of radius r_0 with variations $\hat{\delta r}(k) = \sigma(k)$ at different frequencies k .

Notice that for our parameterized set of curves, the curve derivative in Euclidean coordinates is given as follows (taking into account that $\dot{x}_0^x(t) = \dot{x}_0^y(t) = 0$):

$$\dot{\gamma}(t) = (\delta \dot{r}(t) \cos(t) - (r_0 + \delta r(t)) \sin(t), \delta \dot{r}(t) \sin(t) + (r_0 + \delta r(t)) \cos(t)). \quad (10)$$

The squared absolute value of $\dot{\gamma}(t)$ is then given by

$$|\dot{\gamma}(t)|^2 = \dot{\delta r}(t)^2 + (r_0 + \delta r(t))^2, \quad (11)$$

so that we can write:

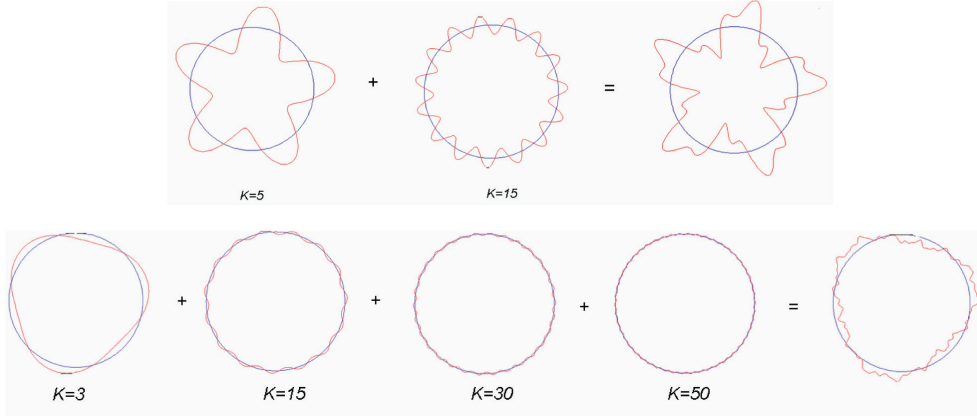
$$E_{\text{smth}}(\gamma) = \int_{[0,2\pi]} dt (\dot{\delta r}(t)^2 + (r_0 + \delta r(t))^2) = 2\pi \left(\sum_{k \in \mathbb{Z}} (k^2 + 1) |\hat{\delta r}(k)|^2 + 2r_0 \hat{\delta r}(0) + r_0^2 \right).$$

Thus, the prior energy E_{curve} takes the form:

$$E_{\text{curve}}(\gamma) = 2\pi \sum (k^2 + 1 + g(k)) |\hat{\delta r}(k)|^2 + 2\pi 2r_0 \hat{\delta r}(0) + 2\pi r_0^2. \quad (12)$$

Defining a function $f(k) = k^2 + 1 + g(k)$, and dropping both $\frac{1}{\pi} r_0^2$, which is simply an additive constant, and the linear term $\frac{1}{\pi} r_0 \hat{\delta r}(0)$, which serves only to change the mean of $\hat{\delta r}(0)$, which we define to be zero, we can write:

Figure 1. Top: shape obtained as a perturbation of a circle at frequencies $k = \{5, 15\}$, with $\hat{\delta}r(5) = 5$, and $\hat{\delta}r(15) = 2$, where $r_0 = 20$. Bottom: shape obtained as a perturbation of a circle at frequencies $k = \{3, 15, 30, 50\}$, with $\hat{\delta}r(3) = 1.5$, $\hat{\delta}r(15) = 0.5$, $\hat{\delta}r(30) = 0.3$, $\hat{\delta}r(50) = 0.2$, where $r_0 = 20$.



$$E_{\text{curve}}(\gamma) = 2\pi \sum_{k \in \mathbb{Z}} f(k) |\hat{\delta}r(k)|^2. \quad (13)$$

Now using the inverse Fourier Transform of $f(k)$ and taking into account equation (9), equation (13) can be written as:

$$\begin{aligned} E_{\text{curve}}(\gamma) &= \\ &= \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} f(k) \iint_{[0, 2\pi]} dt dt' e^{ikt} e^{-ikt'} \delta r(t) \delta r(t') \\ &= \frac{1}{2\pi} \iint_{[0, 2\pi]} dt dt' \delta r(t) \delta r(t') \sum_{k \in \mathbb{Z}} e^{ik(t-t')} f(k) \\ &= \frac{1}{2\pi} \iint_{[0, 2\pi]} dt dt' \delta r(t) \delta r(t') F(t-t'). \end{aligned} \quad (14)$$

Thus we obtain equation (6) where

$$F(t-t') = \sum_{k \in \mathbb{Z}} \exp ik(t-t') f(k). \quad (15)$$

The algorithm makes uses of the functional derivative of E :

$$\frac{\delta E}{\delta \gamma(t)} = \left(\frac{\partial E}{\partial x_0(\gamma(t))}, \frac{\delta E}{\delta(\delta r(\gamma(t)))} \right), \quad (16)$$

where the components taking the following forms:

$$\begin{aligned} \frac{\partial E}{\partial x_0^i(\gamma(t))} &= \int_{[0, 2\pi]} dt (-n(\gamma(t))) \cdot \partial_i \nabla I(\gamma(t)) \\ &+ \int_{[0, 2\pi]} dt (-n(\gamma(t))) \cdot \partial_i U(\gamma(t)), \end{aligned} \quad (17)$$

with U a vector field defined to satisfy $\nabla \cdot U(x) = (G - \bar{G})(x)$ and i indexing Euclidean coordinates on \mathbb{R}^2 ; and

$$\begin{aligned} \frac{\delta E}{\delta(\delta r(t))} &= \lambda_g(r_0 + \delta r(t)) \nabla^2 I(\gamma(t)) \\ &+ \lambda_g(r_0 + \delta r(t)) (G - \bar{G})(\gamma(t)) \\ &+ \frac{1}{\pi} \int_{[0, 2\pi]} dt' F(t-t') \delta r(t'). \end{aligned} \quad (18)$$

Multiple Objects: Model and Algorithm
The multiple-object space is the ‘exponential’ of the single-object space, *i.e.* it consists of all configurations of zero or more objects:

$$\Omega_{\Gamma_o} = \bigcup_{n=0}^{\infty} [\Gamma_o^n / S_n], \quad (19)$$

where S_n indicates the symmetric group of n elements acting on the components of the product. Note that the map τ extends to a map from $\Omega_{\mathbf{C}}$ (the exponential of \mathbf{C}) to Ω_{Γ_o} . Elements of $\Omega_{\mathbf{C}}$ will be denoted ω .

Energy

Given a (real, bounded below) function $H(\omega)$ on $\Omega_{\mathbf{C}}$, we define the Gibbs distribution μ_{β} in terms of the density $p(\omega) = \frac{d\mu_{\beta}}{d\lambda}(\omega)$ w.r.t. Lebesgue-Poisson measure λ on $\Omega_{\mathbf{C}}$:

$$p(\omega) = \frac{z^{|\omega|}}{Z_{\beta}} \exp\{-\beta H(\omega)\}, \quad (20)$$

with parameters $\beta > 0$, $z > 0$ and a normalizing factor Z_{β} :

$$Z_{\beta} = \int_{\Omega_{\mathbf{C}}} d\lambda(\omega) z^{|\omega|} \exp\{-\beta H(\omega)\}. \quad (21)$$

The energy $H(\omega)$ takes the form:

$$H(\omega) = c_0 \sum_i H_1(\omega_i) + \sum_{i \neq j} H_2(\omega_i, \omega_j), \quad (22)$$

where ω_i are the components of ω and c_0 is a weighting parameter. The data term H_1 , is defined as

$$H_1(\omega_i) = E(\tilde{\omega}_i). \quad (23)$$

The term H_2 is the interaction term, which controls the relations between objects, and, in particular discourages overlaps. It is defined as

$$H_2(\omega_i, \omega_j) = \frac{A(R(\tilde{\omega}_i) \cap R(\tilde{\omega}_j))}{\min(A(R(\tilde{\omega}_i)), A(R(\tilde{\omega}_j)))} + \delta_{\varepsilon}(\omega_i, \omega_j), \quad (24)$$

where

$$\delta_{\varepsilon}(\omega_i, \omega_j) = \begin{cases} \infty, & |x_0^i - x_0^j| \leq \varepsilon; \\ 0, & \text{otherwise.} \end{cases} \quad (25)$$

A is the area functional and δ_{ε} is a hard-core repulsion that prevents two components of ω from coinciding (to some tolerance ε). Indeed, if there were no interactions between objects except that they should not coincide, then the optimal configuration of objects would consist of the negative energy elements in Γ_o . The interaction term was thus introduced to prevent the ‘condensation’ of an infinite number of the lowest energy single-object in configuration.

Sampling and Estimation

In order to estimate the configuration of the objects in the image, we use maximum a posteriori estimation, performed by sampling from the probability distribution μ_{β} and applying an annealing scheme. The sampling uses a Markov chain in $\Omega_{\mathbf{C}}$ consisting of a discrete-time multiple birth-and-death process describing all possible transitions from the configuration ω to the configuration $\omega' \cup \omega''$, where $\omega' \subset \omega$ and ω'' is any new configuration. The transition probabilities of this Markov chain take the form:

$$P(\omega \rightarrow \omega' \cup \omega'') = \prod_{\omega'_i \in \omega'} \frac{1}{1 + \delta d_{\beta}} \prod_{\omega''_i \in \omega''} \frac{\delta d_{\beta}}{1 + \delta d_{\beta}} \times (z\delta)^{|\omega''|} \quad (26)$$

where d_{β} is the intensity of the death step of the process detailed below.

This Markov chain can be considered as an approximation of a continuous-time reversible process and converging to it (Descombes,

Minlos, & Zhizhina, 2009), which, within a logarithmic annealing scheme, guarantees uniform convergence to the measure concentrated on the global minima of the energy function $H(\omega)$.

Algorithm Description

We define C as the set of circles lying in the image domain V , with radii in the range $[r_{\min}, r_{\max}]$ and with centres at the image pixels. The curves are represented by a chain of points in \mathbf{R}^2 defined to correspond to discrete parameter values $t_n = 2\pi n / N$ for $n \in \{0, \dots, (N-1)\}$. The circles in C are assumed to have arc length parameterisation, and thus will have equally spaced points.

The birth step of the process adds an unknown number of circles to the current configuration with an intensity z that is independent of the current temperature $T = 1 / \beta$. The death step removes a number of components from the current configuration with a probability that depends on the current (inverse) temperature β and the energy difference $\Delta_i H(\omega) = H(\omega \setminus \omega_i) - H(\omega)$. In more detail, the algorithm is as follows:

Initialization

Discretisation step $\delta = \delta_0$; inverse temperature $\beta = \beta_0$; Poisson mean z_0 ; radius range $[r_{\min}, r_{\max}]$; parameters in E ;

Birth

- (a) Sample a configuration of circles with radii uniformly distributed on $[r_{\min}, r_{\max}]$, from the Lebesgue-Poisson distribution with intensity $z = \delta z_0$, with the addition of a hard core repulsion δ_ε with ε equal to one pixel, producing configuration $\omega \in \Omega_C$;
- (b) Evolve every circle in ω using gradient descent, with gradient field given by equation (16), until convergence, producing

configuration $\tilde{\omega} \in \Omega_{T_0}$; and add the obtained elements to the current configuration (which is empty at the first step);

Death

- (a) For computational efficiency, sort the components of the current configuration w.r.t. their energy $H_1(\omega_i) = E(\tilde{\omega}_i)$;

Remove each component ω_i from the current configuration with probability

$$p_d(\omega_i, \omega) = \frac{\delta d_\beta(\omega_i, \omega)}{1 + \delta d_\beta(\omega_i, \omega)},$$

where

$$d_\beta(\omega_i, \omega) = e^{-\beta \Delta_i H(\omega)}; \quad (27)$$

Termination

If all the components added in the birth step are removed in the following death step, then stop; if not, then decrease the temperature

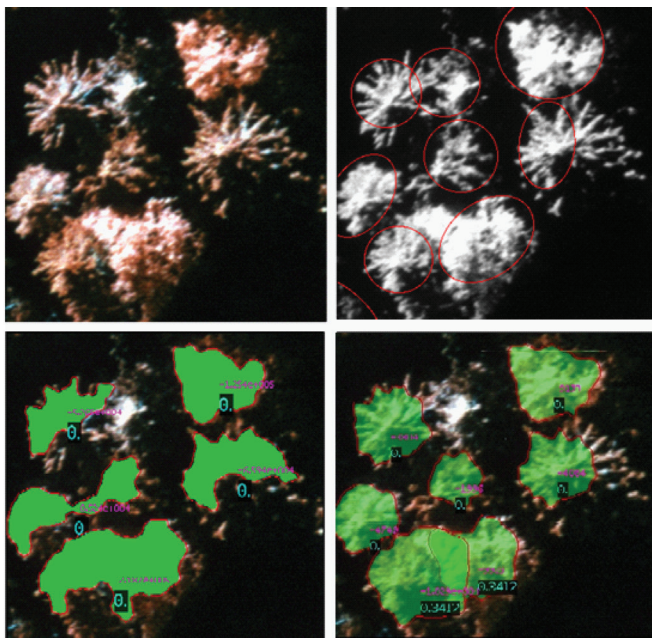
$T = \frac{1}{\beta}$ and time step δ , (we use a geometric annealing schedule to T and δ), and go to the birth step.

Experimental Results

Figures 2-5 show the results of experiments on one band of very high resolution (VHR) colour infra-red (CIR) aerial images. The images show the top part of the tree crowns. The image viewpoint is close to the nadir, *i.e.* the tree crowns are seen from almost vertically above.

Figure 2 demonstrates the results of object extraction obtained using three different models. The top right image shows the configuration obtained using simply-shaped objects, in this case ellipses. The bottom left image shows the configuration obtained using the representation and energy described in this paper, but with no

Figure 2. Top left: CIR image of tree crowns, © CBA. Top right: final configuration using ellipse-shaped objects. Bottom left: final configuration using arbitrarily-shaped objects obtained using an MPP without shape prior. Bottom right: final configuration of arbitrarily-shaped objects obtained using an MPP with shape prior. The numbers show the interaction term value H_2 (black background) and the data term value H_1 .



strong shape information, *i.e.* with $E_{sh} \equiv 0$, or equivalently, $g \equiv 0$. The bottom right image shows the result obtained using the full energy described in this paper, *i.e.* with strong prior shape information included.

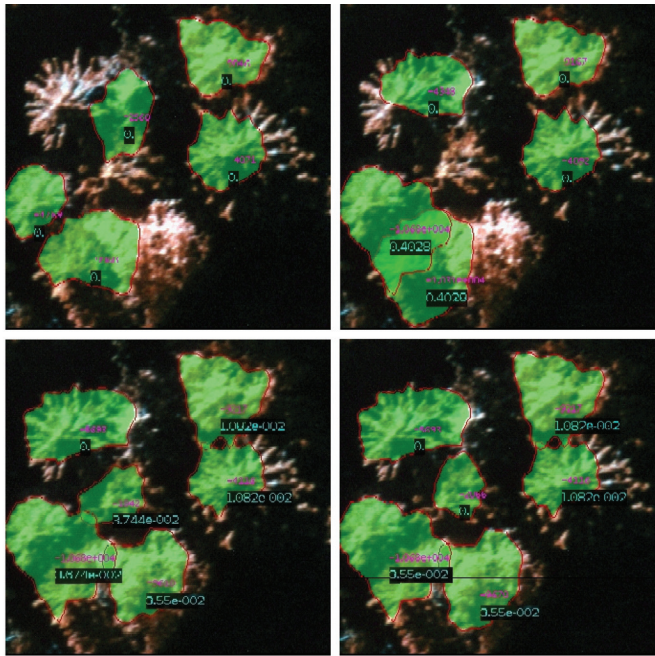
Note that, for all three experiments, there is no curve initialization, *i.e.* the initial configuration is empty. The first birth step creates a certain number of curves depending only on the Poisson mean z_0 , but the final number is determined automatically by the convergence of the annealed birth-and-death process.

As we have said, in the images used for our experiments, the tree crowns are pictured from vertically above. They thus have shapes that are roughly circular with perturbations corresponding to the tree branches or leaves. For the third experiment, the function $g(k)$ is

defined so as to discourage low frequencies, and in particular $k = 2$, in order to prevent the extraction of several nearby crowns as elongated objects, as well as to favour roughly circular objects with small perturbations corresponding to branches or leaves. The strong shape information allows us to extract nearby and even overlapping objects without paying the price of greatly simplifying their geometry, in contrast to classical Marked Point Process models based on low-dimensional parametric objects such as ellipses.

To further demonstrate the performance of the extended MPP model including strong prior shape knowledge, we present another result confirming the discussion of the previous paragraph. We did an experiment where the objects are defined so that elongation is no longer penalized, *i.e.* the frequency $k = 2$ is

Figure 3. Evolution of the configuration during the birth-and-death process. The objects are obtained using the energy E , where $g(k)$ is defined so as to discourage low frequency perturbations, except the frequency $k = 2$ which is relaxed. Bottom right: final configuration obtained. The numbers on the crowns show the interaction term value H_2 (black background) and the data term value H_1 .



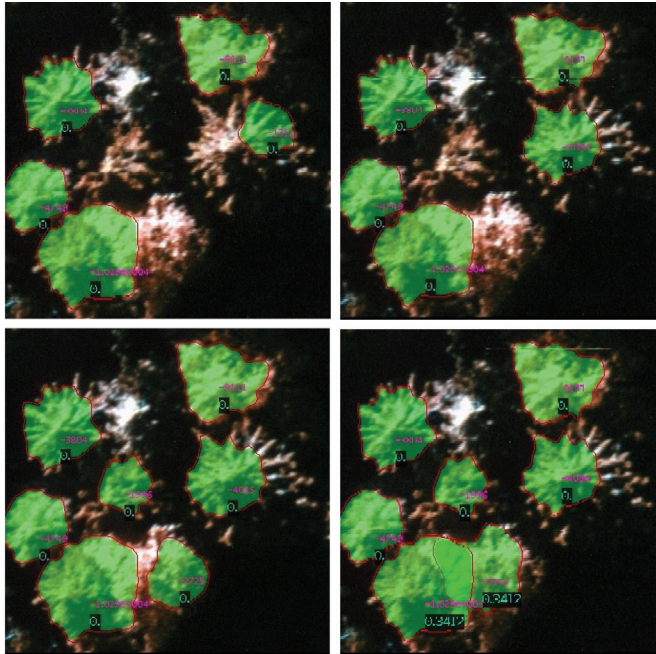
relaxed. Figure 3 shows the evolution of a configuration composed of such objects, whereas Figure 4 illustrates the evolution of the configuration when elongation is discouraged. From the results, we can see that when $g(k)$ is defined so as to favour $k = 2$, trees are extracted as elongated objects, contrary to the case when $k = 2$ is penalized. This can also lead to the extraction of two nearby trees as a single elongated object; *cf.* the group of trees in the bottom left part of the image.

Finally, Figure 5 shows the result of an experiment on another, bigger image of 900×900 pixels using the same model parameters as for the third experiment in Figure 2. The image represents a complex scene containing many overlapping, visible as well as obscured tree crowns of various sizes, but they

are separated into distinct objects thus allowing delineation of the individual crowns. The computation time of the C++ programme was 85 minutes on a 2.16 GHz processor.

To summarize, an MPP model using simply shaped marks allows the rapid detection of the objects in an image, but the geometrical accuracy is very low for objects with complex shapes. Therefore, this type of model is an appropriate tool for the detection of objects in high resolution images. The second approach, an MPP model for the extraction of arbitrarily-shaped objects with only weak shape information (Kulikova, Descombes, Jermyn, Zerubia, & Zhizhina, 2010), is geometrically far more accurate, while not increasing the computational complexity unduly. The limitations of this approach, however, are that because it uses only

Figure 4. Evolution of the configuration during the birth-and-death process. The objects are defined using the energy E , where $g(k)$ is defined so as to discourage low frequency perturbations, and in particular the frequency $k = 2$. Bottom right: final configuration obtained. The numbers on the crowns show the interaction term value H_2 (black background) and the data term value H_1 .



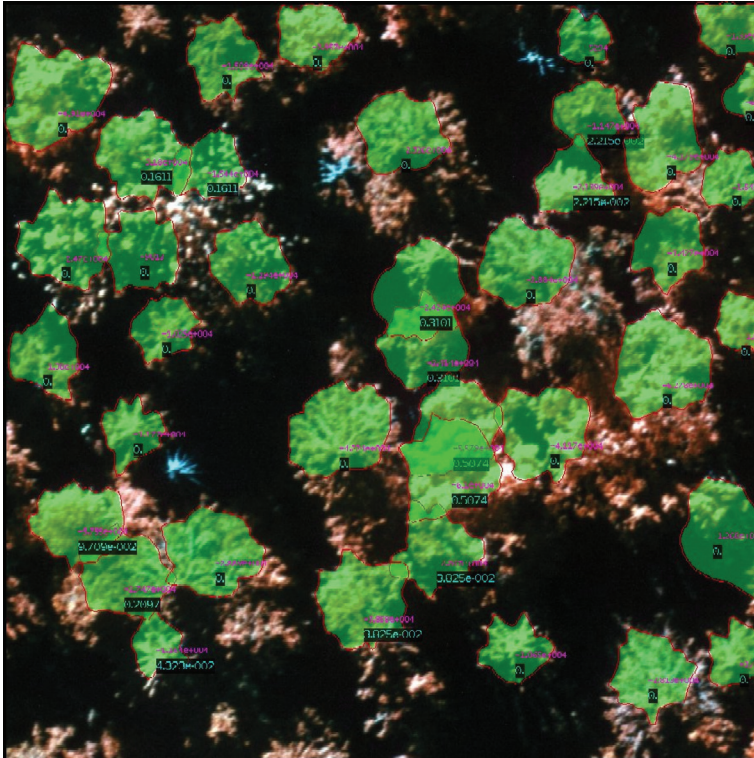
weak shape information to define the possible single objects, it cannot, first, detect different types of objects with similar radiometric characteristics; and second, separate two or more overlapping objects. The MPP model for arbitrarily-shaped objects including strong prior shape information described in this paper deals with both these limitations. For example, in the bottom left corner of the bottom left image Figure 2, there are two overlapping tree crowns that are extracted as one object using the second approach, but which are extracted as two distinct overlapping objects using the model presented in this paper. Using strong prior shape information has another advantage: even if the object to be extracted is partly obscured or cluttered, by, for example, shadow, like the spiky crowns of spruce trees, the object may

nevertheless, to some degree of accuracy, be correctly extracted as a single object.

CONCLUSION AND PERSPECTIVES

The marked point process framework with simple geometries that has been used in the past for the extraction of objects from images has been recently extended to arbitrarily shaped objects. In this paper, we extend this approach further by incorporating strong prior knowledge about the shape of the objects sought, without increasing the dimensionality of the single-object space (and thereby the computational complexity). The set of possible single objects is defined using the local minima of an energy that incorporates information coming from the

Figure 5. Final configuration obtained (original CIR image of tree crowns, © CBA). The numbers on the crowns show the interaction term value H_2 (black background) and the data term value H_1 .



data and strong prior shape information about the objects sought. This allows us to deal with configurations of overlapping objects with complex shapes.

The birth-and-death algorithm used to minimize the multiple-object energy over configurations of multiple objects has the advantage that, at every iteration, the current configuration is updated by adding multiple objects independently of the current energy and temperature, contrary to the classical RJMCMC algorithm, which increases the computational efficiency.

The work in this paper can also be seen as the first step in a joint tree extraction and classification algorithm. For the class of images used in our experiments, once the tree crowns are extracted, one can then classify

them into species (Kulikova, Mani, Srivastava, Descombes, & Zerubia, 2007).

The next step consists in further extending the model to the extraction of objects of several classes in scenes of high complexity containing overlapping, arbitrarily-shaped objects.

ACKNOWLEDGMENTS

This work was partly conducted within the INRIA/IITP/UIIP Associated Team 'ODESSA' (<http://www-sop.inria.fr/ariana/Projets/Odessa/>) and the INRIA/FSU Associated Team 'SHAPES' (<http://www-sop.inria.fr/ariana/Projets/Shapes/>). We are thankful to the Swedish University of Agricultural Sciences of Uppsala for providing us with data.

REFERENCES

- Caselles, V., Kimmel, R., & Sapiro, G. (1997). Geodesic active contours. *International Journal of Computer Vision*, 22(1), 61–79. doi:10.1023/A:1007979827043
- Cremers, D., Osher, S. J., & Soatto, S. (2006b). Kernel density estimation space and intrinsic alignment for shape priors in level set segmentation. *International Journal of Computer Vision*, 69(3), 335–351. doi:10.1007/s11263-006-7533-5
- Cremers, D., Sochen, N., & Schnörr, C. (2006a). A multiphase dynamic labelling model for variational recognition-driven image segmentation. *International Journal of Computer Vision*, 66(1), 67–81. doi:10.1007/s11263-005-3676-z
- Cremers, D., Tischauser, F., Weickert, J., & Schoerr, C. (2002). Diffusion snakes: Introducing statistical shape knowledge into the Mumford-Shah functional. *International Journal of Computer Vision*, 50(3), 295–313. doi:10.1023/A:1020826424915
- Descamps, S., Descombes, X., Béchet, A., & Zerubia, J. (2009). Détection de flamants roses par processus ponctuels marqués pour l'estimation de la taille des populations. *Traitement du Signal*, 28(2).
- Descombes, X., Minlos, R., & Zhizhina, E. (2009). Object extraction using a stochastic birth-and-death dynamics in continuum. *Journal of Mathematical Imaging and Vision*, 33(3), 347–359. doi:10.1007/s10851-008-0117-y
- Joshi, S. H., & Srivastava, A. (2009). Intrinsic Bayesian active contour for extraction of object boundaries in images. *International Journal of Computer Vision*, 81, 331–355. doi:10.1007/s11263-008-0179-8
- Juan, O., Keriven, R., & Postelnicu, G. (2006). Stochastic motion and the level set method in computer vision: Stochastic active contours. *International Journal of Computer Vision*, 69(1), 7–25. doi:10.1007/s11263-006-6849-5
- Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contours models. *International Journal of Computer Vision*, 1(4), 321–331. doi:10.1007/BF00133570
- Kulikova, M., Jermyn, I., Descombes, X., Zerubia, J., & Zhizhina, E. (2010). Extraction of arbitrarily-shaped objects using stochastic multiple birth-and-death dynamics and active contours. In *Proceedings of the SPIE Electronic Image Processing Conference*.
- Kulikova, M., Mani, M., Srivastava, A., Descombes, X., & Zerubia, J. (2007). Tree species classification using radiometry, texture and shape based features. In *Proceedings of the European Signal Processing Conference*.
- Lacoste, C., Descombes, X., & Zerubia, J. (2005). Point processes for unsupervised line network extraction in remote sensing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10), 1568–1579. doi:10.1109/TPAMI.2005.206
- Leventon, M. E., Grimson, W. E. L., & Fougères, O. (2000). Statistical shape influence in geodesic active contours. *Computer Vision and Pattern Recognition*, 1, 316–323.
- Ortner, M., Descombes, X., & Zerubia, J. (2007). Building outline extraction from digital elevation models using marked point processes. *International Journal of Computer Vision*, 72(2), 107–132. doi:10.1007/s11263-005-5033-7
- Osher, S., & Fedkiw, R. (2003). *Level set methods and dynamic implicit surfaces*. Berlin, Germany: Springer-Verlag.
- Perrin, G., Descombes, X., & Zerubia, J. (2005). A marked point process model for tree crown extraction in plantation. In *Proceedings of the IEEE International Conference on Image Processing* (pp. 661–664).
- Sethian, J. A. (1999). *Level set methods and fast marching methods*. Cambridge, UK: Cambridge University Press.
- Srivastava, A., & Jermyn, I. H. (2009). Looking for shapes in two-dimensional cluttered point clouds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(9), 1616–1629. doi:10.1109/TPAMI.2008.223
- Storvik, G. (1994). A Bayesian approach to dynamic contours through stochastic sampling and simulated annealing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(10), 976–986. doi:10.1109/34.329011
- Tu, Z., & Zhu, S.-C. (2002). Image segmentation by data-driven Markov Chain Monte Carlo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5), 657–673. doi:10.1109/34.1000239