Shape recognition for image scene analysis

Maria Kulikova

UNIVERSITY OF NICE - SOPHIA ANTIPOLIS - FRANCE

GRADUATE SCHOOL STIC
INFORMATION AND COMMUNICATION TECHNOLOGIES AND SCIENCES

THESIS

to fulfill the requirements for the degree of

DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE

from the University of Nice - Sophia Antipolis

Specialized in : CONTROL, SIGNAL AND IMAGE PROCESSING

presented by

Maria S. KULIKOVA

SHAPE RECOGNITION FOR IMAGE SCENE ANALYSIS

Supervised by Ian H. JERMYN, Xavier DESCOMBES, and Josiane ZERUBIA and prepared at
INRIA Sophia-Antipolis Méditerranée in the ARIANA research team

Defended the 16^{th} of December, 2009

Jury:

M. Albert BIJAOUI	Astronomer, 1^{st} class, OCA	President
M. Albert DIPANDA	Professor, University of Bourgogne	Reviewer
M. Zoltan KATO	Associate Professor, University of Szeged	Reviewer
M. Anuj SRIVASTAVA	Professor, Florida State University	Examiner
Mrs. Elena ZHIZHINA	Professor, IITP of Moscow	Examiner
Mrs. Josiane ZERUBIA	Director of Research, INRIA	Director
M. Xavier DESCOMBES	Director of Research, INRIA	Supervisor
M. Ian H. JERMYN	Senior Permanent Researcher, INRIA	Supervisor

Acknowledgements

During my doctoral study, I had the chance to meet many pleasant and interesting people. These acknowledgements are to thank all these people for their friendship, help and encouragement.

First of all, this work would not have taken place without the support of my director Josiane Zerubia and my supervisors Ian H. Jermyn and Xavier Descombes. I would like to thank them for having confidence in me and for giving me the opportunity to work on shape modelling in relation to image analysis.

I would like to thank Professors Elena Zhizhina, Albert Bijaoui, Anuj Srivastava, Zoltan Kato, and Albert Dipanda for accepting to be members of my thesis jury and for their interest and kind attention during the defence. I am also very grateful to Anuj and his PhD student Shantanu for the enriching discussions I had with them and for helping me to achieve a deeper understanding of shape analysis both during their stays at INRIA and my visit to FSU. Then, I would like to thank Elena for clarifying several aspects of the Point Process methodology and for her contagious happiness whether at INRIA, at IITP or during our walking and visiting. Finally, thank you to Elena and Anuj for the fruitful collaborations from which much of the work of this thesis was produced.

I had the honour to work with Professor Robert Adolfovitch Minlos, whom I would like to thank for his interest in my work and for giving me numerous new ideas as a result of our discussions during his stays in the Ariana project.

Thanks to my colleagues in Ariana and its neighbourhood: Caro, Gui, Pierrot, Alexis, Saloua, Meena, Floflo, Alex and Sap, Peter, Olivier, Stig, Praveen and Leïla, Rafa, Aymen, Dani, Gio, Ahmed, Bayrem, Carlito, Csaba, Vova, Sylvain, Mikael, Eric ... and our project assistants Corinne and Laurie, for their help, relaxing company and just listening, as well as for the funny moments spent together during the Ariana activities and festivities. I would particularly like to thank two people who welcomed me very warmly on my arrival in Ariana and helped me to adapt to this new lab; thank you for their attention and friendship, and finally for their advice and encouragement at the key moments throughout my stay in the project when I needed it the most: a thousand thanks, Gui and Pierrot.

Thanks a lot to Lena, Elaine, Christophe, Alban, and Guillaume, for their support on the

day of my defence and also to all those who were present.

My special thanks go to Ian. I thank him very much for his interest and his great availability: hours of enriching discussions which made me discover an enormous amount within the universe of computer science and beyond it. Thanks to him for all this; especially during the last year of my thesis, the most interesting and important, and thus the heaviest year, for his encouragement and precious advice which helped me to avoid going round in circles and to finally go straight to the point.

I am infinitely grateful to my family for always supporting and encouraging me; to my mother Galina, my brother Ilya, and my grand parents Polina Mikhajlovna and Viktor Vasilievitch, who wanted me to graduate so much, which sometimes gave me additional strength.

Finally, my greatest thanks are to Colin for all his critical comments and advice, understanding, psychotherapy and encouragement ... and for those so delicious cakes that he cooked for the after-defence party. But, of course, and especially, for always being there for me over these last three years.

Biot, March, 2010

Contents

General introduction

There are many of domains, *e.g.* medicine, the military, biology, astronomy, and remote sensing, where a huge variety of problems is treated by making use of image analysis techniques such as image filtering, image fusion, image synthesis, image measurements, and many others. In medicine, for example, images (or sequences) of the human body and organs, are used for examination purposes or the detection of anomalies. In astronomy, there is a special discipline, called astrography, where astronomical objects are observed and studied using photographs. In remote sensing, image data is used for Earth observation and related problems. The data is collected using different types of device, installed on aircraft or satellites. Remote sensing allows us to collect data from inaccessible areas, as well as to replace the high cost and slow data acquisition by humans on the ground, for monitoring, for example, changes in glaciers or forested zones.

In this thesis, we focus on image analysis applied to remote sensing problems. We address two main topics: 'Tree species study and classification', presented in Part I, and 'Multiple arbitrarily-shaped object extraction with marked point processes', described in Part II. These parts are further introduced below.

Remote sensing aerial images are now widely used by the forestry organisations in order to facilitate their observations on the ground. The data that can be extracted from images of forest provide important information for forest monitoring, *e.g.* the age of the trees and thus the average age of the forest, the density of trees in plantations, or the diversity of tree species in the forest. For the last purpose the classification of tree crowns visible in the image into species is a crucial problem. To tackle this problem, numerous techniques using radiometric as well as textural tree characteristics have been developed (several of them will be listed in the introduction to the first part of this manuscript). We propose to take into account information about the shapes of tree crowns in order to improve on these classification schemes. The use of this information is now pertinent since:

- The resolution of optical satellite and aerial images is continually increasing. Nowadays, resolutions range from several tens of centimetres down to several centimetres. At these resolutions, the geometry of objects at the human scale is clearly visible;
- The human visual cortex remains the best apparatus for image analysis. Often, to recognize, and thus to detect or extract an object, it is sufficient for a person to look at the silhouette of the object. Thus, due to the knowledge acquired during his life, the person is able to recognize the object knowing only its shape.

Thus, inevitably, geometry has to be taken into account in image analysis.

In Part I, therefore, we study the classification of tree crowns into species using shape information, without, or in combination with, radiometric and textural features. The shapes analysed consist of boundaries of tree crowns extracted from very high resolution aerial infra-red images.

In order to classify the trees, their crowns should first be detected and segmented from images. The second topic we address in this dissertation is **the arbitrarily-shaped multiple** object extraction problem from very high resolution remote sensing images using a stochastic approach based on object processes, the Marked Point Processes (MPP), which are well known for their ability to include geometrical information about objects. In previous work, MPP models have been used for the extraction of objects from images of lower resolution where the objects had simple geometrical shapes and were thus represented using only simply-shaped objects, *e.g.* rectangles, discs, or ellipses. Our aim is to lift this restriction, *i.e.* to define a marked process model in the space of arbitrarily-shaped objects. In Part II, thus, we first present an MPP model that includes weak shape information about object boundaries, and then, we propose a method for incorporating strong prior shape knowledge into the model.

Dans beaucoup de domaines comme la médecine, la surveillance, la biologie, l'astronomie et la télédétection, les techniques d'analyse d'image comme le filtrage d'image, la fusion d'images et beaucoup d'autres, sont fréquement utilisées. En médecine par exemple, des anomalies sont détectées en examinant des images du corps humain. En astronomie, une discipline, appelée astrographie, consiste à observer et à étudier les objets en utilisant leurs photographies. En télédétection, les données image sont utilisées dans un but d'observation de la Terre. La télédétection via des installations embarqueées à bord d'avion ou de satellites, permet à la fois de collecter des informations sur des endroits inaccessibles et d'éviter les contraintes et les coûts élevés de l'acquisition d'information par des humains au sol (par exemple pour observer les changements des glaciers ou les zones forestières).

Cette thèse traite de l'analyse d'image appliquée aux problématiques de télédétection et se décompose en deux thèmes majeurs : 'Etude et classification d'espèces d'arbres' et 'Extraction de multiples objets de formes arbitraires à l'aide de processus ponctuels marqués' décrits respectivement dans les parties I et II.

La télédétection sur les images aériennes est d'ores et déjà largement utilisée par les organisations forestières pour faciliter leurs tâches et les observations sur le terrain. Il est possible d'extraire des informations importantes des images pour l'observation des forêts comme par exemple : l'âge des arbres et donc l'âge moyen de la forêt, la densité des arbres dans les plantations, ou encore la diversité des espèces composant la forêt. Pour ce faire, de nombreuse techniques utilisant les caractéristiques radiométriques ou de texture des arbres ont déjà été développées (certaines d'entre elles sont listées dans l'introduction de la partie I du manuscrit). L'approche de cette thèse est de prendre en consideration l'information de la forme des couronnes des arbres (*i.e.* les houppiers) pour améliorer les shémas de classification. L'utilisation de cette information est justifiée par les raisons suivantes:

- La résolution des images issues des satellites d'observation et des images aeriennes augmente continuellement ; aujourd'hui elle varie de quelques dizaines de centimètres à quelques centimètres. A de telles résolutions, la géométrie des objets est clairement visible;
- Le cortex visuel humain est encore le meilleur 'outil' d'analyse d'image. Pour reconnaitre un objet, et donc le détecter et l'extraire, il est souvent suffisant pour une personne de regarder la silhouette de l'objet. Grâce à la connaissance acquise au cours de leur vie, les humains sont capables de reconnaitre les objets par leur seule forme.

Par conséquent, la géométrie est un critère pertinent à prendre en compte pour l'analyse d'image.

Dans la partie I, nous étudions donc la classification des arbres en function de leur couronne en utilisant l'information sur leur forme; sans ou en combinant cette information avec des critères radiométrique et de texture. Les formes analysées sont les contours des couronnes

des arbres extraits d'images infra-rouge couleur (IRC) de très haute résolution.

Le second thème traité dans ce manuscrit est l'extraction d'objets multiples de formes arbitraires dans des images de très haute résolution par l'utilisation d'une approche stochastique, fondée sur des processus objet : les Processus Ponctuels Marqués (PPM). Ces processus sont connus pour la possibilité qu'ils offrent d'inclure des informations géométriques sur les objets. Dans des travaux antérieurs, les PPMs ont été utilisés pour l'extraction d'objets dans des images de moindre résolution, où les objets avaient des formes géométriques simples et étaient donc représentés par des formes simples : rectangles, disques, ellipses. Le travail développé dans cette seconde partie de thèse consiste à lever cette restriction de simplicité de la forme, c'est-à-dire à définir un modèle de PPM utilisant un espace de forme arbitraire d'objets. La partie II présente donc un premier modèle de PPM qui incorpore une information a priori faible de forme sur les contours des objets, puis une méthode d'incorporation d'une information a priori forte des formes recherchées dans le modèle.

Part I

Tree species classification

Chapter 1

Introduction

Interest in applying remote sensing to forest studies goes back to the 1920s when aerial photographs were first used to assess forest inventory. Remote sensing is now widely employed in forest management where the aerial information is combined with measurements taken on the ground to study the biodiversity of the forest ecosystem. The methods developed in forest image processing aim to facilitate the task of forest inventory and assessment.

The most useful parameters obtained from aerial images and ground measurements are density of planting, age of trees, stem volume, tree species composition, and information about biotopes and habitats that have ecological value. Some of these parameters can be estimated by interpretation of aerial photographs. However, manual interpretation takes time and requires costly human resources when compared to the use of automatic techniques.

In order to obtain information about, for example, the diversity of forest species or stem volume, the classification of tree crowns into species is necessary. Prior to the classification step, the segmentation of the image into individual tree crowns is necessary. Segmentation techniques such as template matching, edge detection and others that have been employed for this application are discussed in [Erikson, 2004b], [Perrin *et al.*, 2006], and will be studied in the second part of this dissertation.

A few approaches have been proposed to classify segmented tree crowns into species. One method is the *Signature Generation Process* where for every crown extracted, a class of signatures is created from the multi-spectral data of the initial image, (cf [Leckie *et al.*, 2003]). A likelihood maximisation technique is used to label the crowns. Some crowns with signatures too far from a successful match remain unclassified. In another study [Gougeon, 1995a], tree crowns are manually delineated to avoid bias due to bad detection. 50 trees for each class verified against the ground truth are selected. Seven parameters, such as the multi-

spectral average of pixels in the crown (average on each of the bands) and the illuminated part of the crown, as well as the multi-spectral value of the brightest pixel in the tree crown are then computed. Spectral signatures of a crown or a region within a crown are developed by combining mean and covariance patterns. The distinct characteristics allow us to regroup delineated crowns in forested populations [Gougeon, 1996]. There are some limitations with this method due to the close spectral signature of species like the red cedar and the fir, for instance [Gougeon *et al.*, 1998]. In [Erikson, 2004a], a classification based on reflectance is used to separate the conifers from the deciduous trees. The internal structure and shading within a crown offer other differentiating criteria. One such measure, the proportion of red and white pixels to the total number of pixels, can identify birch trees. Another helps to differentiate spruce trees using the concavities of the crown shape created by the shadow of the tree top. A hierarchy of criteria is set forth to classify the crowns. Classification accuracy using this strategy was 75%.

Radiometry and texture analysis have been used extensively in remote sensing applications. We propose to incorporate information obtained from studying **tree crown shapes** to improve classification performance. The shapes are closed planar curves representing the boundaries of tree crowns extracted from very high resolution Colour InfraRed (CIR) aerial images. Figure 1.1 shows one such image.

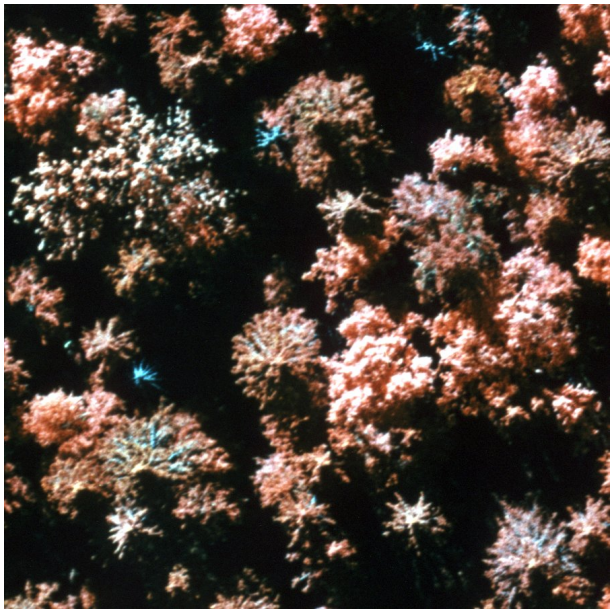

Figure 1.1: An example of one of the images used in this work (3cm resolution), provided by the Swedish University of Agricultural Science, © CBA, Uppsala.

The images from which the tree crowns are extracted, are provided by the Centre of Image Analysis (CBA) of Uppsala (Sweden). They are Colour InfraRed (CIR) images of 3cm/pixel resolution, representing forested zones where trees are close to the nadir, *i.e.* the tree crowns are seen from almost vertically above. The CIR film provides more information about vegetation than natural colours. In fact, it allows deciduous trees to be distinguished from conifers and healthy trees from diseased trees, because it is very dependent on chlorophyll activity. Once digitized, the CIR photographs are represented by artificial colours to describe the three bands [NIR, Red, Green]: the red in the image represents NIR; the green represents the red; and the blue represents the green.

The classification consists in assigning a species to each extracted tree, (cf Fig. 1.2).

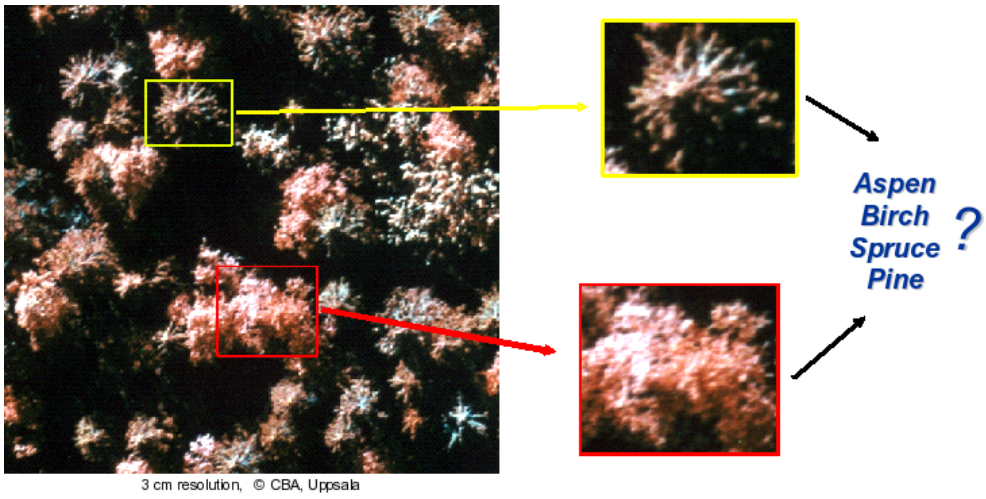

Figure 1.2: Original image is the CIR image of 3cm resolution, © CBA, Uppsala.

As in the Erikson [Erikson, 2004a] study, tree classification is performed on the four most prevalent species in Sweden: Norway spruce, Scots pine, birch, and aspen. Two of these species are coniferous while the other two are deciduous. We select 48 crowns (12 per class, (cf Fig. 1.3)). Their contours, represented by a cyclically ordered set of points, are then extracted in order to study their shapes, (cf Fig. 1.4).

The support vector machine (SVM) [Vapnik, 1995], a supervised learning method, was chosen for the classification. An important property of this classifier is that during the training process, only a small subset of the training set vectors are selected as support vectors. This reduces computational cost and provides better generalization, so that, for instance, when new samples far from the decision boundary are introduced, the existing support vectors remain unchanged. The main points of the approach are described in Appendix A.

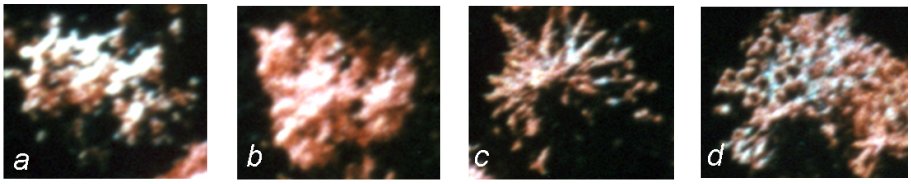

Figure 1.3: Examples of the four most prevalent species in Sweden: (a) Aspen, (b) Birch, (c) Norway spruce, (d) Scots pine.

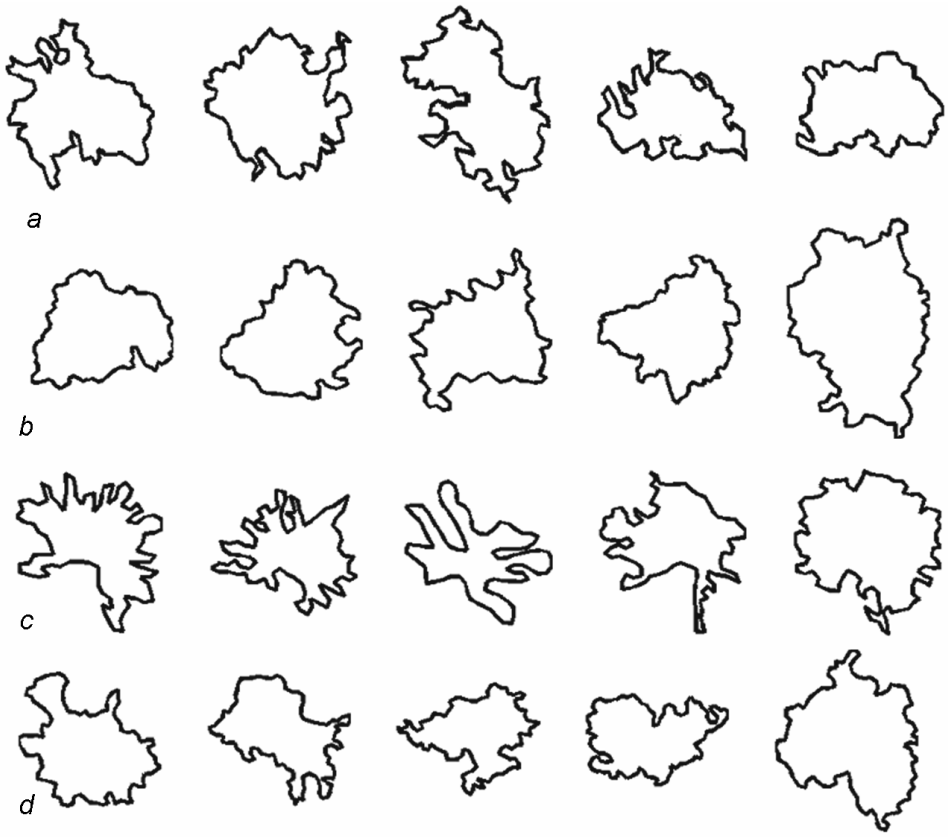

Figure 1.4: Examples of tree crown shapes: (a) Aspen, (b) Birch, (c) Norway spruce, (d) Scots pine.

This work was partly performed in collaboration with M. Mani, who made a study of the radiometry and texture attributes of tree crowns and created an ensemble of features based on information obtained [Kulikova *et al.*, 2007]; this will be detailed in chapter 3.

The reminder of this part is organized as follows. In chapter 2, we present a study of tree crown shape under different shape representations, comparing the two methodologies on several applications. In chapter 3, we describe the features considered and we present the classification results obtained: the idea is to study the shapes of the tree crowns in the four classes and develop a set of features which will be combined with some conventional radiometry and texture based features to see if this improves the results of classification using only radiometry and texture information.

Chapter 2

Studies of tree crown shape using 'shape spaces'

In this chapter, we present a study of tree crown shapes under different representations. We consider both the angle-function representation and the elastic q-representation in the corresponding shape spaces, and use geodesic distance in shape space to compare shapes or compute features. Firstly, the computation of geodesics is presented in section 2.1, and several examples of geodesics are illustrated. Then, in section 2.2, the two approaches will be compared from point of view of tree crown classification.

2.1 Geodesic distance as a measure of the similarity of tree crown shapes

For our study, we chose a methodology based on the shape analysis of closed continuous curves on shape spaces using geodesic paths. We first perform studies using the bending metric with the angle-function curve representation developed by Klassen *et al.* [Klassen *et al.*, 2004], and then using the elastic metric and the square root q-function representation proposed by Joshi *et al.* [Joshi *et al.*, 2007a, Joshi *et al.*, 2007b]. This methodology was chosen for several reasons. The shapes of tree crowns extracted from images of 3cm/pixel resolution have a complex structure. The hypothesis of continuity avoids the choice of reference points (landmarks) on the curve. Most existing approaches in the domain of shape analysis suffer from restrictions due to the use of reference points describing the shapes, *e.g.* how to choose these points or how to extract them. 'Shape spaces' with appropriate metrics are then defined as non-linear manifolds using the shape invariance properties. This framework allows the formulation of statistical inferences on such shape manifolds. Finally, with every geodesic path in shape space is associated a distance. We tested this distance to see if it can be used as a measure of the similarity of tree crown shapes in the sense of belonging to one or another species.

2.1.1 Bending metric: angle-function shape representation

The first metric proposed for the study of tree crown shape is the bending metric [Klassen *et al.*, 2004]. Here, we recall the main points of the framework.

In this framework, shape characterization is general in the following sense:

- Every curve, representing a shape, is a continuous curve;
- Shapes are elements of a space, called shape space.

The main idea of shape analysis in this framework is to use the differential geometry of curves and of the space of curves. This analysis consists of two steps:

1. Differential geometric representation of shapes;
2. Development of algorithms for computing geodesic paths between arbitrary shapes in the shape space.

Tree crown shape representation

We consider tree crown contours as continuous and closed curves in $\mathbb{R}^2$. The representation of such curves in the shape space is invariant to rigid rotation and translation, and to scaling in $\mathbb{R}^2$, called shape preserving curve transformations. Let us define these properties.

Curves $\alpha = (\alpha_1(s), \alpha_2(s))$ are parametrized by arc-length s, where $\alpha : \mathbb{R} \to \mathbb{R}^2$ with period 2π satisfying the condition of constant speed along the curve: $|\alpha'(s)| = 1$, $\forall s$. We can write $\alpha'(s) = e^{j\theta(s)}$, where $\theta : \mathbb{R} \to \mathbb{R}$ and $j = \sqrt{-1}$ associating $\mathbb{C}$ with $\mathbb{R}^2$. θ is called the *direction function* or *angle function*. For every s, $\theta(s)$ gives the angle between $\alpha'(s)$ and the positive x axis, (cf Fig. 2.1).

Figure 2.2 shows the angle-functions for a circle and an ellipse. Graphs of more complex angle functions representing tree crowns are presented in figure 2.3.

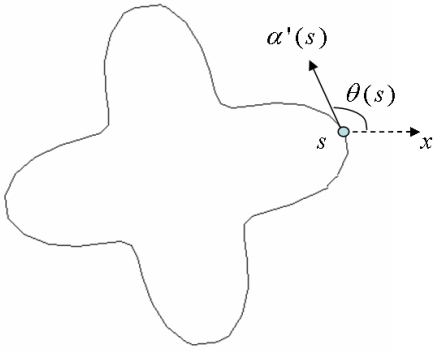

Figure 2.1: Illustration of the definition of the angle function.

We denote by $\mathbb{L}^2$ the space of real, 2π periodic functions $\mathbb{R} \to \mathbb{R}$, square integrable on $[0, 2\pi]$ with scalar product $\langle f_1, f_2 \rangle = \int_0^{2\pi} f_1(s) f_2(s) ds$ and $\|f\| = \sqrt{\langle f, f \rangle}$.

Let $\theta(s)$ be the angle function of a planar shape. For the unit circle this function is $\theta_0(s) = s$. For any other closed curve the angle function can be written as $\theta = \theta_0 + f$, where $f \in \mathbb{L}^2$. The properties of curve invariance mentioned above are guaranteed by the following conditions (cf [Klassen *et al.*, 2004]):

- The problem of scaling can be simply resolved by fixing the length of the curve, for instance to be 2π. Let $\bar{c}$ be a closed contour, then we obtain:

$$\int_{\bar{c}} ds = \int_0^{2\pi} ds = 2\pi;$$

- To guarantee the invariance to translation, the curve must satisfy the condition:

$$\int_0^{2\pi} \alpha(s) ds = 0,$$

  i.e. the curve's centre of 'mass' is at origin;

- Adding a constant to the angle function θ is equivalent to a rotation of the curve in $\mathbb{R}^2$. To guarantee the invariance of the curve to this action, we deal with the functions θ, the mean values of which are equal to a constant on $[0, 2\pi]$. So, let $\frac{1}{2\pi} \int_0^{2\pi} \theta(s) ds = \pi$, where constant π is chosen to include the identity function θ_0 in the restricted set, since

$$\frac{1}{2\pi} \int_0^{2\pi} \theta_0(s) ds = \frac{1}{2\pi} \int_0^{2\pi} s ds = \pi;$$

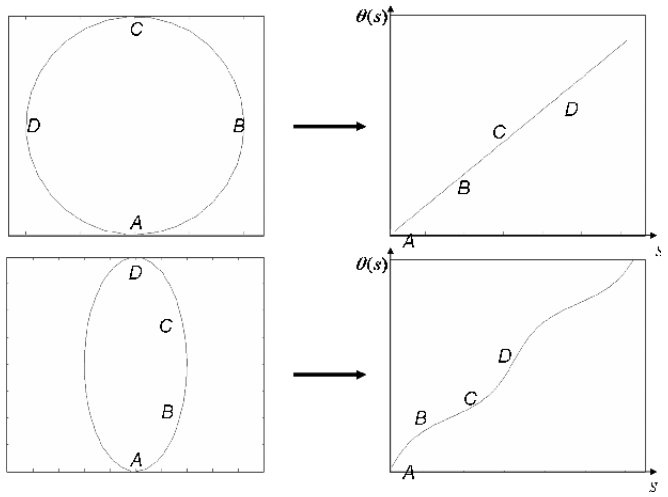

Figure 2.2: Examples of shapes and the corresponding angle functions.

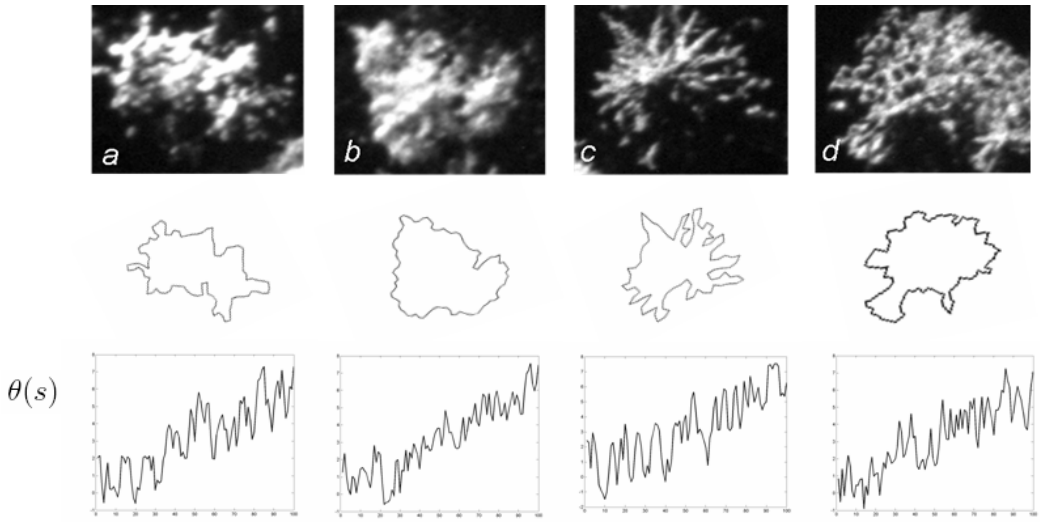

Figure 2.3: Examples of crowns of four species, their contours, and the corresponding angle functions θ, where (a) stands for aspen, (b) for birch, (c) for spruce, and (d) for pine.

- Finally, to be closed, the curves must satisfy following condition:

$$\int_0^{2\pi} \exp(j\theta(s))ds = 0.$$

We define $\mathcal{C} \subset \theta_0 + \mathbb{L}^2$ as the set of all the elements of $\theta_0 + \mathbb{L}^2$ satisfying the conditions

described above. Or more formally, define the map $\phi = (\phi^1, \phi^2, \phi^3)) : (\theta_0 + \mathbb{L}^2) \to \mathbb{R}^3$:

$$\phi^1(\theta(s)) = \frac{1}{2\pi} \int_0^{2\pi} \theta(s) ds$$

$$\phi^2(\theta(s)) = \int_0^{2\pi} \cos(\theta(s)) ds$$

$$\phi^3(\theta(s)) = \int_0^{2\pi} \sin(\theta(s)) ds.$$

From now on, we can define $\mathcal{C}$ as $\mathcal{C} = \phi^{-1}(\pi, 0, 0)$, which is called the pre-shape space, because the same curve with different initial points ($s = 0$) corresponds to different elements of the space $\mathcal{C}$.

As the geometry of $\mathcal{C}$ is too complex to allow an analytic expression for geodesics on this space, [Klassen *et al.*, 2004] suggested approximating geodesics on $\mathcal{C}$ by building up a series of infinitesimal lines in the space $\theta_0 + \mathbb{L}^2$ containing the pre-shape space $\mathcal{C}$, and then projecting them back to $\mathcal{C}$. To do that, it is necessary to define a mechanism of projection from the space $\theta_0 + \mathbb{L}^2$ to $\mathcal{C}$.

Pre-shape space geometry

Tangent and normal subspaces of the pre-shape space

To facilitate the construction of the tangent space, first, the space of normals in $\mathcal{C}$ is constructed with the help of the map ϕ used for the definition of $\mathcal{C}$. The derivative $d\phi$ at point $\theta \in \theta_0 + \mathbb{L}^2$ and in direction $f \in \mathbb{L}^2$ of the map $\phi : \theta_0 + \mathbb{L}^2 \to \mathbb{R}^3$ is given by the following equations:

$$d\phi^1(f) = \frac{1}{2\pi} \int_0^{2\pi} f(s) ds = \left\langle f, \frac{1}{2\pi} \right\rangle$$

$$d\phi^2(f) = -\int_0^{2\pi} \sin(\theta(s)) f(s) ds = -\langle f, \sin(\theta) \rangle \tag{2.1}$$

$$d\phi^3(f) = \int_0^{2\pi} \cos(\theta(s)) f(s) ds = \langle f, \cos(\theta) \rangle$$

This implies that the vector $f \in \mathbb{L}^2$ is tangent to the space $\mathcal{C}$ at θ if and only if f is orthogonal to the sub-space spanned by $\{1, \cos(\theta), \sin(\theta)\}$. The normal space at θ is therefore defined by these three functions. The tangent space is then a space of co-dimension three defined in the following way:

$$T_\theta(\mathcal{C}) = \left\{ f \in \mathbb{L}^2 | f \perp \mathbf{span}\{1, \cos(\theta), \sin(\theta)\} \right\}$$

Projection on the pre-shape space

To construct the geodesic on the space $\mathcal{C}$, the projection of a point in $\mathbb{L}^2$ on the space $\mathcal{C}$ is necessary. For this purpose, the notion of a level set of the map ϕ is used. The idea is to move in the direction perpendicular to the level set such that the image by ϕ of the path between a point and its projection to $\mathcal{C}$ is a segment in $\mathbb{R}^3$ (cf Fig. 2.4).

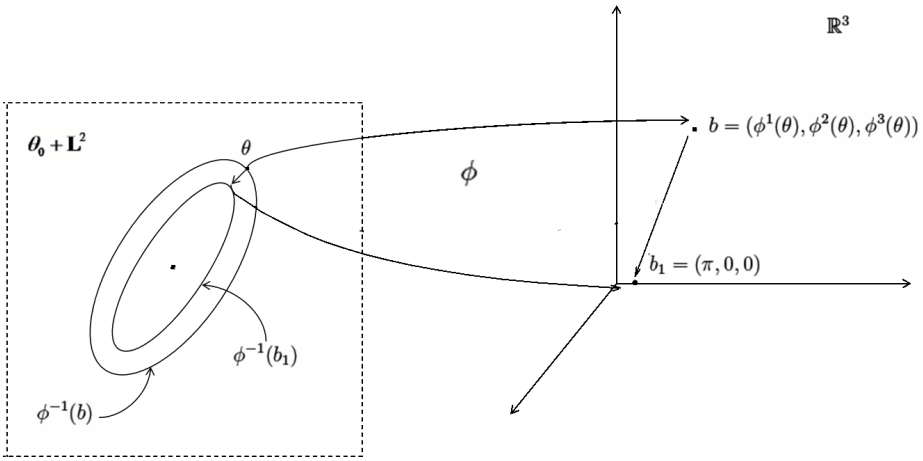

Figure 2.4: Illustration of the projection process of a point θ of $\mathbb{L}^2$ to the space $\mathcal{C}$.

Consider the set $\phi^{-1}(b) = \left\{ \theta \in \theta_0 + \mathbb{L}^2 | \phi(\theta) = b \right\}$ for any point $b \in \mathbb{R}^3$. The level set for $b_1 = (\pi, 0, 0)$ represents the pre-space $\mathcal{C}$. Given a point $\theta \in \phi^{-1}(b)$, we define a shift $d\theta$ along the direction orthogonal to the level set, bringing us thus to $\mathcal{C}$. As ϕ is a map from $\mathbb{L}^2$ to $\mathbb{R}^3$, the Jacobian $d\phi$ is a map from the tangent space $T_\theta(\mathbb{L}^2)$ to the tangent space $T_b(\mathbb{R}^3)$.

Define $d\theta$ to be the normal vector at θ such that $\phi(\theta + d\theta) = b$. The Jacobian matrix J is then given by

$$J = \begin{bmatrix} \left\langle \frac{1}{2\pi}, 1 \right\rangle & \left\langle \frac{1}{2\pi}, \sin(\theta) \right\rangle & \left\langle \frac{1}{2\pi}, \cos(\theta) \right\rangle \\ \langle -\sin(\theta), 1 \rangle & \langle -\sin(\theta), \sin(\theta) \rangle & \langle -\sin(\theta), \cos(\theta) \rangle \\ \langle \cos(\theta), 1 \rangle & \langle \cos(\theta), \sin(\theta) \rangle & \langle \cos(\theta), \cos(\theta) \rangle \end{bmatrix}. \tag{2.2}$$

Let $r(\theta) = b_1 - \phi(\theta)$ be the residue. Now, the projection procedure from $\theta \in \mathbb{L}^2$ to $\mathcal{C}$ is given by algorithm 2.1. We denote this projection $\prod : \mathbb{L}^2 \mapsto \mathcal{C}$.

Algorithm 2.1 Projection $\prod(\theta)$ of $\theta \in \mathbb{L}^2$ to $\mathcal{C}$,

1. Initialize $\epsilon > 0$;
2. Compute the residue $r(\theta) = b_1 - \phi(\theta)$;
3. If $\|r(\theta)\| < \epsilon$, then stop; if not, then continue;
4. Calculate the Jacobian $J(\theta)$ defined by (2.2);
5. Compute $d\theta = \beta_1 + \beta_2 \sin(\theta) + \beta_3 \cos(\theta)$, where $\beta = J(\theta)^{-1} r(\theta)$;
6. Update the curve: $\theta = \theta + d\theta$ and go to step 2;

Geodesics on pre-shape space

As mentioned above, to construct a geodesic on the pre-shape space $\mathcal{C}$ the idea is to approximate it by successive small increments along the tangent direction in $\theta_0 + \mathbb{L}^2$ and to project them to $\mathcal{C}$. In detail, the algorithm is as follows. We are given $\theta \in \mathcal{C}$ and $f \in T_\theta(\mathcal{C})$ the tangent vector at θ. The geodesic starting at θ in the direction f is generated by the one-parameter flow denoted $\Psi(\theta, t, f)$, where t is a time parameter. The flow is calculated for the discrete time $t = \Delta, 2\Delta, 3\Delta, \ldots\ 1$ for a small value of $\Delta > 0$. Initialize $\Psi(\theta, 0, f) = \theta$ for $t = 0$. Then, add a first small Δ to θ in the direction f to reach the new point $\theta + \Delta f \in \mathbb{L}^2$ and project it to $\mathcal{C}$ using $\prod$ and set $\Psi(\theta, \Delta, f) = \prod(\theta + \Delta f)$ giving the next point of the geodesic. This process is repeated until $t = 1$. It remains to 'transport' the tangent vector f to the new obtained point on the geodesic in way that preserves its length and its angle w.r.t. the geodesic according to the definition of geodesic, cf Def. 2.1. Let $\tilde{\theta}$ be a new point of the geodesic where a vector $\tilde{f}$ is tangent to $\mathcal{C}$ at $\tilde{\theta}$ and parallel to f. $\tilde{f}$ is then given by

$$\tilde{f} = \|f\| \frac{g}{\|g\|},\ g = f - \sum_{k=1}^{3} \langle f, h_k \rangle h_k, \tag{2.3}$$

where $\{h_k\}$ forms an orthonormal basis of the normal space: $\mathbf{span}\left\{1, \cos(\tilde{\theta}), \sin(\tilde{\theta})\right\}$.

Definition 2.1. *A **geodesic** on a manifold embedded in a Euclidean space is a curve whose tangent vector remains parallel if it is transported along it, which means that the acceleration vector is perpendicular to the tangent plane to the surface at every point along the geodesic curve.*

Finally, the algorithm for computing geodesics is given by algorithm 2.2. This algorithm provides a set of points approximating the geodesic path between the initial shape represented by θ and the point reached at $t = 1$.

Algorithm 2.2 Construction of a geodesic starting at θ in the direction f

1. Initialize: $\theta \in \mathcal{C}$, direction $f \in T_\theta(\mathcal{C})$, $l = 0$, $\Psi(\theta, l\Delta, f) = \theta$, and choose $\Delta > 0$;
2. Compute $\Psi(\theta, l\Delta, f) + \Delta f$ as $\Psi(\theta, (l+1)\Delta, f) = \prod(\Psi(\theta, l\Delta, f) + \Delta f)$;
3. Transport f obtained in the previous iteration to new point $\tilde{\theta}$ using equation (2.3);.
4. Put $l = l + 1$ and $f = \tilde{f}$;
5. If $l\Delta = 1$, then stop; if not, go to step 2.

Geodesics on shape space

So far, the process of geodesic path construction has been defined in the pre-shape space $\mathcal{C}$ where different curves could represent identical shapes differing only in their starting points $s = 0$. A shift of the starting point along the curve defines an action of the group $\mathbb{S}^1 = \mathbb{R}/2\pi\mathbb{Z}$ on $\mathcal{C}$. The space of curves invariant to starting point shifts is given by the pre-shape space modulo the unit circle $\mathbb{S}^1$: $\mathcal{S} = \mathcal{C}/\mathbb{S}^1$ and called shape space. The algorithm for geodesic construction remains the same, except, that in (2.3), the $\{h_k\}$ now form the orthonormal basis generated by $\mathbf{span}\left\{1, \cos(\tilde{\theta}), \sin(\tilde{\theta}), \dot{\tilde{\theta}}\right\}$, where $\dot{\tilde{\theta}}$ is a derivative of $\tilde{\theta}$.

Numerical method for geodesic computation

Algorithm 2.2 describes how to compute a geodesic path along the given direction f starting at some initial point θ. To compute now the shortest path that relates a starting point $\theta_1 \in \mathcal{S}$ and $\theta_2 \in \mathcal{S}$, it is necessary to find a direction $f \in T_{\theta_1}(\mathcal{S})$ such that the flow along this direction reaches θ_2 at the moment $t = 1$: $\Psi(\theta_1, 1, f) = \theta_2$. The solution is given by the minimum over f of the functional $H(f) = \|\Psi(\theta_1, 1, f) - \theta_2\|^2_{\mathbb{L}^2}$.

To solve this problem, since $f \in \mathbb{L}^2$, a finite dimensional approximation of f by its Fourier coefficients is used: $f(s) \approx \sum_{n=0}^{m}(a_n \cos(ns) + b_n \sin(ns))$ for m big enough. The functional $H(f)$ takes then the following form:

$$\tilde{H}(a, b) = \left\| \Psi(\theta_1, 1, \sum_{n=0}^{m} a_n \cos(ns) + b_n \sin(ns)) - \theta_2 \right\|^2_{\mathbb{L}^2} .$$

Gradient descent is then applied to find the optimal solution.

Figure 2.5 shows several examples of the transformation of one curve into another along a geodesic path (every curve represents a tree crown boundary).

There is a distance associated with every path. The distance is obtained by the summation of the projections approximating the geodesic path. Figures 2.6, 2.7, 2.8 show the

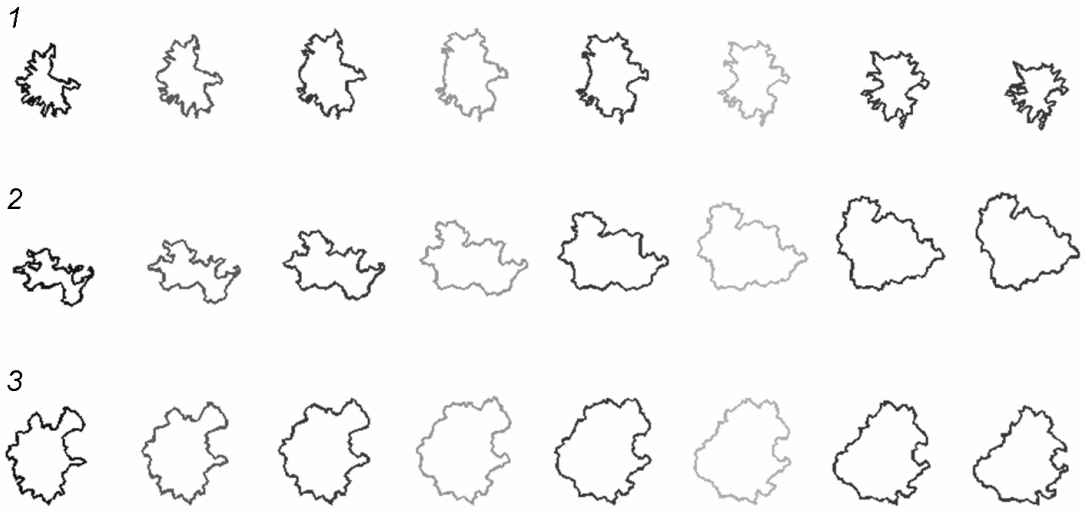

Figure 2.5: Examples of geodesic paths for trees of different species using the angle-function representation.

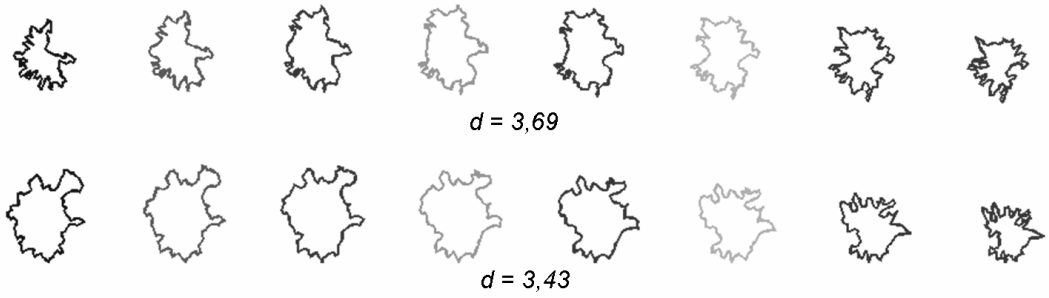

Figure 2.6: Top: geodesic between two spruce trees. Bottom: geodesic between a pine and a spruce.

geodesics and the corresponding distances d for a tree of one species and another tree of the same species, and the same tree from the first species and a tree from different species. The results show that the similarity of two trees in the sense of geodesic distance using the bending metric does not mean that they belong to the same species, and, therefore, cannot be used as is for classifying trees into species.

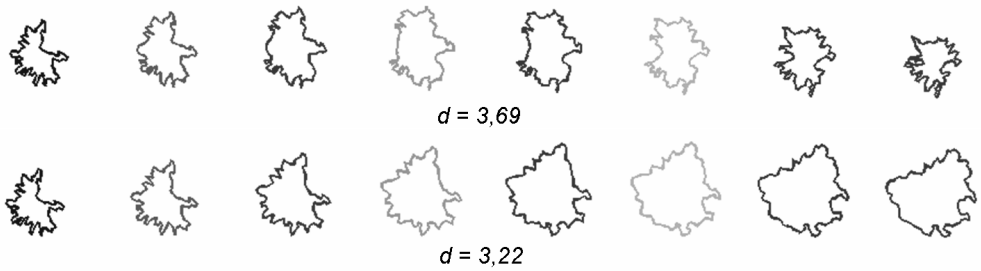

Figure 2.7: Top: geodesic between two spruces. Bottom: geodesic between a spruce and a birch.

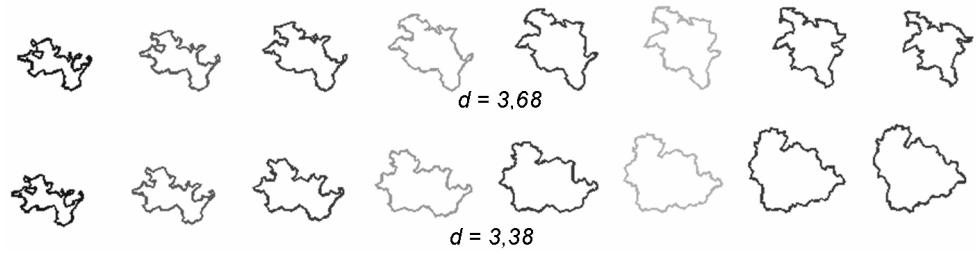

Figure 2.8: Top: geodesic between two aspen trees. Bottom: geodesic between an aspen and a birch.

2.1.2 Elastic metric: q-function shape representation

The second metric used for the study of tree crown shape, is based on the Riemannian analysis of elastic continuous closed curves in the *'q-function'* representation, proposed by Joshi *et al.* and detailed in [Joshi *et al.*, 2007a], [Joshi *et al.*, 2007b], and [Joshi, 2007].

The main point of this approach is that it combines an elastic metric for shape analysis and a path-straightening approach to finding geodesics. The elastic metric originally proposed by Younes [Younes, 1998], allows a curve to stretch as well as to bend as it deforms along a geodesic in the corresponding shape space. The path-straightening approach, in contrast to the shooting method for geodesic construction described in the previous section, uses the idea of initializing a geodesic by an arbitrary path, and then successively evolves (straightens) it under the gradient of the geodesic energy.

Curve representation

We consider tree crown boundaries represented by closed planar curves $\gamma : [0, 2\pi] \to \mathbb{R}^2$, being an element of $\mathbb{L}^2_1([0, 2\pi])$, meaning that its first derivative exists and that both $\gamma(t)$ and $\gamma'(t)$ lie in $\mathbb{L}^2([0, 2\pi])$.

In the shape space, an elastic curve is represented by a function $q(t) : [0, 2\pi] \rightarrow \mathbb{R}^2$ as follows,

$$q(t) = \frac{\dot{\gamma}(t)}{\sqrt{||\dot{\gamma}(t)||}} \quad , \tag{2.4}$$

where $||\cdot||$ is the standard Euclidean $\mathbb{R}^2$ norm. $q(t)$ is an infinite-dimensional vector-function in $\mathbb{L}^2(\mathbb{R}^2)$. Let $\mathcal{Q}$ be the set of such elements.

Physically, $||\dot{\gamma}(t)||$ can be viewed as the square-root of the instantaneous speed along the curve, and $\frac{\dot{\gamma}(s)}{||\dot{\gamma}(t)||}$ represents the instantaneous direction along the curve. Within a translation the curve is then given by $\gamma(s) = \int_0^s dt\, q(t)\ ||q(t)||$.

Curve invariance

In the same way as for the angle function representation, the algorithm for geodesic construction is first defined on the pre-shape space.

The representation of curves in the shape space is invariant to rigid translations and uniform scaling.

- The problem of scaling is solved by fixing the length of the curve, which is given by:

$$\int_{[0,2\pi]} dt\ ||q(t)||^2 = 1.$$

  Then, the space of curves invariant to uniform scaling and rigid translation is defined by:

$$\mathcal{B} \equiv \{q : [0, 2\pi] \rightarrow \mathbb{R}^2 | \int_{[0,2\pi]} dt\ ||q(t)||^2 = 1\} \quad . \tag{2.5}$$

  The tangent space of $\mathcal{B}$ is then given as:

$$T_q(\mathcal{B}) = \{w = (w^1, w^2) | w : [0, 2\pi] \rightarrow \mathbb{R}^2 | \int_{[0,2\pi]} dt\, (w(t), q(t))_{\mathbb{R}^2} = 0\} \quad . \tag{2.6}$$

  The space $\mathcal{B}$ is an infinite dimensional unit sphere, on which the geodesics are great circles and can be defined analytically. Therefore, the geodesic between two elements $q_0, q_1 \in \mathcal{B}$ along a unit direction $f \in T_{q_0}(\mathcal{B})$ towards q_1 for time t can be calculated using the formula:

$$\begin{aligned} \chi_t(q_0, f) &= \cos\left(t \arccos \int_{[0,2\pi]} ds\, (q_0, q_1)_{\mathbb{R}^2}\right) q_0 \\ &\quad + \sin\left(t \arccos \int_{[0,2\pi]} ds\, (q_0, q_1)_{\mathbb{R}^2}\right) f \quad . \end{aligned} \tag{2.7}$$

To parallel transport a tangent vector f from any point q_0 on the sphere to q_1, the map $\pi : T_{q_0}(\mathcal{B}) \to T_{q_1}(\mathcal{B})$ is defined as follows:

$$\pi(f; q_0, q_1) = f - 2\frac{(q_0 + q_1)\int_{[0,2\pi]} ds\,(f, q_1)_{\mathbb{R}^2}}{\int_{[0,2\pi]} ds\,(q_0 + q_1, q_0 + q_1)_{\mathbb{R}^2}} \quad . \tag{2.8}$$

- To be closed, the curves in $\mathcal{B}$ must satisfy the following condition:

$$\int_{[0,2\pi]} dt\,\dot{\gamma}(t) = \int_{[0,2\pi]} dt\,q(t)\;\|q(t)\| = 0 \quad . \tag{2.9}$$

To define the space of the elements satisfying the closure condition, consider a map $\mathcal{G} \equiv (\mathcal{G}^1, \mathcal{G}^2)$ as:

$$\begin{aligned} \mathcal{G}^1 &= \int_{[0,2\pi]} dt\,q^1(t)\;\|q(t)\|\,, \\ \mathcal{G}^2 &= \int_{[0,2\pi]} dt\,q^2(t)\;\|q(t)\| \quad . \end{aligned}$$

The space of closed curves is given then by $\mathcal{A} = \mathcal{G}^{-1}(0,0)$.

The space of curves satisfying both conditions is then as $\mathcal{C} = \mathcal{A} \cap \mathcal{B} \subset \mathcal{Q}$ and is called the pre-shape space of elastic arbitrary-speed parametrized curves.

Pre-shape space geometry

Tangent and normal spaces to the pre-shape space

To define the tangent space to the pre-shape space $\mathcal{C}$ at point q, first, the normal space at q is defined. Given tangent vectors $u, v \in T_q(\mathcal{Q})$ at q, the scalar product on $T_q(\mathcal{Q})$ is defined by:

$$\langle u, v\rangle = \int_{[0,2\pi]} dt\,(u(t), v(t))_{\mathbb{R}^2} \quad . \tag{2.10}$$

The derivative $d\mathcal{G}$ at point q and in direction $w \in T_q(\mathcal{Q})$ is:

$$\begin{aligned} d\mathcal{G}^1(w(t)) &= \int_{[0,2\pi]} dt\,(w(t), \frac{q^1(t)}{\|q(t)\|}q(t) + \|q(t)\|\,\mathbf{e}^1)_{\mathbb{R}^2} = \left\langle w, \frac{q^1(t)}{\|q(t)\|}q(t) + \|q(t)\|)_{\mathbb{R}^2}\mathbf{e}^{\mathbf{1}}\right\rangle \\ d\mathcal{G}^2(w(t)) &= \int_{[0,2\pi]} dt\,(w(t), \frac{q^2(t)}{\|q(t)\|}q(t) + \|q(t)\|\,\mathbf{e}^2)_{\mathbb{R}^2} = \left\langle w, \frac{q^2(t)}{\|q(t)\|}q(t) + \|q(t)\|_{\mathbb{R}^2}\,\mathbf{e}^2\right\rangle \quad , \end{aligned}$$

where $\mathbf{e}^i$, $i = 1, 2$ are the identity matrix i^{th} columns. The normal space to $\mathcal{A}$ is then defined as follows:

$$N_q(\mathcal{A}) = \mathbf{span}\left\{\nabla\mathcal{G}^1(t) = \frac{q^1(t)}{\|q(t)\|}q(t) + \|q(t)\|\,\mathbf{e}^1, \nabla\mathcal{G}^2(t) = \frac{q^2(t)}{\|q(t)\|}q(t) + \|q(t)\|\,\mathbf{e}^2\right\} \quad . \tag{2.11}$$

This implies that a vector $w \in T_q(\mathcal{Q})$ is tangent to $\mathcal{C}$ at point q if it is in $T_q(\mathcal{B})$ and orthogonal to $N_q(\mathcal{A})$, the tangent space to $\mathcal{C}$ taking the following form:

$$T_q(\mathcal{C}) = \{w : [0, 2\pi] \to \mathbb{R}^2 \mid w \in T_q(\mathcal{B}), w \perp N_q(\mathcal{A})\} \quad . \tag{2.12}$$

Projection to the pre-shape space

To construct geodesics on the pre-shape space, a projection of an element to it from the embedding space $\mathcal{Q}$ is necessary. To perform this, the idea is to move in the direction orthogonal to the level set of $\mathcal{G}$ to reach the nearest point in $\mathcal{C}$. In detail, the algorithm is as follows. Consider a residual vector $l(q) = -\mathcal{G}(q)$, $l \in \mathbb{R}^2$ and a shift dq along the direction normal to the level set of $\mathcal{G}$. The Jacobian $J_{i,j} = \left\langle \nabla\mathcal{G}^i, \nabla\mathcal{G}^j \right\rangle$, where $i, j = 1, 2$ mapping $T_q(\mathcal{Q})$ to $T_{\mathbf{0}}(\mathbb{R}^2)$ is as follows:

$$J_{i,j} = \begin{cases} 3\int_{[0,2\pi]} dt\; q^i(t)q^j(t), & i \neq j, \\ 3\int_{[0,2\pi]} dt\; (q^i(t)q^j(t) + 1), & i = j. \end{cases} \tag{2.13}$$

The process of projection is then performed by algorithm 2.3.

Algorithm 2.3 Projection of $q \in \mathcal{Q}$ on $\mathcal{C}$

1. Initialize $\epsilon > 0$;
2. Compute $\|l(q)\|$; if $\|l(q)\| < \epsilon$, then stop; if not, then continue;
3. Compute the Jacobian J given by expression (2.13);
4. Calculate $\beta^T = J^{-1}(q)l^T(q)$;
5. Compute a shift vector $dq = \sum_{i=1,2}\beta^i\nabla\mathcal{G}^i(q)$;
6. Update the curve as $q = q + \delta dq$, for some small $\delta > 0$;
7. Rescale the curve length by $q = \frac{q}{\sqrt{<q,q>_q}}$ and go to step 2;

To project a tangent vector $w \in T_q(\mathcal{Q})$ to $T_q(\mathcal{C})$, the tangent space of $\mathcal{C}$, the vector is first projected to the tangent space of unit-length curves, which is given by

$$\tilde{w} \equiv w - < w, q > q \quad , \tag{2.14}$$

and then it is projected to the tangent space of $\mathcal{C}$ by:

$$w_{proj} \equiv \tilde{w} - \left\langle \tilde{w}, \mathbf{e}^1_{\mathcal{G}(q)} \right\rangle \mathbf{e}^1_{\mathcal{G}(q)} - \left\langle \tilde{w}, \mathbf{e}^2_{\mathcal{G}(q)} \right\rangle \mathbf{e}^2_{\mathcal{G}(q)} \quad , \tag{2.15}$$

where $\left\{\mathbf{e}^i_{\mathcal{G}(q)}\right\}$ is an orthonormal basis of $\left\{\nabla \mathcal{G}^i(q)\right\}$ w.r.t. the scalar product given by equation (2.10).

Geodesics on pre-shape space

To construct a geodesic on the pre-shape space, a variational method is used, which consists in the iterative straightening of an arbitrary path connecting two elements until it becomes a geodesic. Let $\alpha : [0, 1] \to \mathcal{C}$ be a path connecting any two elements $q_0, q_1 \in \mathcal{C}$, and consider the energy functional

$$E[\alpha] = \frac{1}{2} \int_{[0,2\pi]} dt \ \langle \dot{\alpha}(t), \dot{\alpha}(t) \rangle \quad . \tag{2.16}$$

Given an initial path α, we can perform gradient descent in the space of paths between q_0 and q_1 on $\mathcal{C}$ to arrive at a minimum of E corresponding to a geodesic in $\mathcal{C}$. In order to derive the gradient of the energy functional, define $\mathcal{F}$ as the set of all paths in $\mathcal{C}$ and $\mathcal{F}_0 \subset \mathcal{F}$ as the set of paths connecting q_0 and q_1. The tangent space to $\mathcal{F}$ at α is given by $T_\alpha(\mathcal{F}) = \{w \,|w(t) \ \in T_{\alpha(t)}(\mathcal{C}), \ \forall t \in [0, 1]\}$, where the inner product is defined by:

$$\langle\langle u_1, u_2 \rangle\rangle = \langle u_1(0), u_2(0) \rangle + \int_0^1 dt \ \left\langle \frac{Du_1}{dt}(t), \frac{Du_2}{dt}(t) \right\rangle \quad . \tag{2.17}$$

The tangent space to $\mathcal{F}_0$ is given by $T_\alpha(\mathcal{F}_0) = \{w(t) \in T_{\alpha(t)}(\mathcal{F}) \mid w(0) = w(1) = 0\}$, where $w(t)$ is a tangent vector field at point $\alpha(t)$ on the path α.

Theorem 2.1. *The gradient vector field of E in $T_\alpha(\mathcal{F})$ is given by a vector field v such that $\frac{Dv}{dt} = \dot{\alpha}$ and $v(0) = 0$.*

$\frac{Dv}{dt}$ is the covariant derivative of the vector field $v \in T_\alpha(\mathcal{F})$, which is defined as the orthogonal projection of $\frac{dv}{dt}$ on $T_{\alpha(t)}(\mathcal{C})$ for $\forall t$ and for $\alpha \in \mathcal{C}$. The vector field v, which gives the gradient of E, is called the covariant integral of a vector field $w \in T_\alpha(\mathcal{F})$ along α and is given by the vector field $u \in T_\alpha(\mathcal{F})$ such that $\frac{Du}{dt} = w$.

To construct a geodesic, consider a number k of points approximating it. The computation of a path derivative is then given by algorithm 2.4.

For computing a geodesic, parallel transport is used, which is defined in the following way. Let $w_0 \in T_{\alpha(0)}(\mathcal{C})$ be a vector field for $\alpha : [0, 1] \to \mathcal{C}$. Then, there exists a unique parallel vector field $w(t)$ such that $\frac{Dw(t)}{dt} = 0$ and $w(0) = w_0$. $w(t)$ is the forward parallel transport of w_0 along the path α at time t ($\tilde{w}(t) = w(1-t)$ for backward parallel transport). Algorithm 2.5

Algorithm 2.4 Computation of $\frac{d\alpha}{dt}$

1. Initialize $\frac{d\alpha}{dt}(0) = 0$, k and $\tau = 1$;
2. Compute $\theta = \arccos\left\langle \alpha(\frac{\tau-1}{k}), \alpha(\frac{\tau}{k}) \right\rangle$;
3. Calculate the direction $f = -\alpha(\frac{\tau-1}{k}) + \alpha(\frac{\tau}{k})\cos(\theta)$;
4. Compute the velocity vector at the points on the path, which is then given by $\frac{d\alpha}{d\tau}(\frac{\tau}{k}) = \frac{kf\theta}{\sqrt{\langle f,f\rangle}}$;
5. Project $\frac{d\alpha}{d\tau}(\frac{\tau}{k})$ to $T_{\alpha(\frac{\tau}{k})}(\mathcal{C})$ using equations (2.14) and (2.15);
6. Rescale the curve length by $q = \frac{q}{\sqrt{<q,q>_q}}$;
7. Increment $\tau = \tau + 1$; if $\tau > k$ stop; if not, go to step 2;

gives the procedure for the parallel transport of a tangent vector field $w \in T_{\alpha_\tau}(\mathcal{F})$ to the next point α_τ on the geodesic giving a parallel tangent vector field $w^{\|} \in T_{\alpha_{\tau+1}}(\mathcal{F})$ and denoted $w^{\|} = \mathbb{P}\left(w(\frac{\tau-1}{k}); \alpha(\frac{\tau-1}{k}), \alpha(\frac{\tau}{k})\right)$.

Algorithm 2.5 Parallel transport of tangent vector w

1. Initialize $l_w = \langle w, w\rangle$;
2. Compute $w^{\|}$ as a transported w using the expression given by (2.8): $w^{\|} = \pi\left(w(\frac{\tau-1}{k}); \alpha(\frac{\tau-1}{k}), \alpha(\frac{\tau}{k})\right)$;
3. Project $w^{\|}$ into $T_{\alpha(\frac{\tau}{k})}(\mathcal{C})$ using equations (2.14) and (2.15);
4. Rescale the length: $w^{\|} = l_w \frac{w^{\|}}{\left\langle w^{\|}, w^{\|}\right\rangle}$;

The covariant integral of $\dot{\alpha}$ at every point of α can be now computed using the approach described by algorithm 2.6.

The geodesic in $\mathcal{C}$ is given by a path on which the covariant derivative $\frac{D}{dt}\left(\frac{d\alpha}{dt}\right)$ is zero for every $t \in [0, 1]$.

Lemma 2.1. *The orthogonal complement of the tangent space $\mathcal{F}_0$ is given by* $T^{\perp}_{\alpha}(\mathcal{F}_0) \equiv \left\{w \in T_\alpha(\mathcal{F}) \mid \frac{D}{dt}(\frac{Dw}{dt}) = 0\right\}$

Algorithm 2.6 Covariant integration of $\dot{\alpha}$

1. Initialize $w(0)$ to zero, $\tau = 1$ and a number k of points ;
2. Compute $w(\frac{\tau}{k}) = \mathbb{P}\left(w(\frac{\tau-1}{k}); \alpha(\frac{\tau-1}{k}), \alpha(\frac{\tau}{k})\right) + \frac{1}{k}\frac{d\alpha}{dt}(\frac{\tau}{k})$;
3. Project $w(\frac{\tau}{k})$ to $T_{\alpha(\frac{\tau}{k})}(\mathcal{C})$ using equations (2.14) and (2.15);
4. Increment $\tau = \tau + 1$; if $\tau > k$ stop; if not, go to step 2;

From the lemma, it follows that the projection $v \in T_\alpha(\mathcal{F}_0)$ of a tangent field $w \in T_\alpha(\mathcal{F})$ is given by substracting a covariant linear vector field $t\tilde{w}(t)$, where $\tilde{w}(t)$ is the backward parallel transport of $w(1)$ along the path α, *i.e.* for $\forall t \in [0, 1]$

$$u(t) = w(t) - t\tilde{w}(t) \quad . \tag{2.18}$$

The field thus obtained represents the vector gradient field of the energy functional $E[\alpha]$ (cf Fig. 2.9).

The procedure for updating the path α under the energy field gradient is described in algorithm 2.7.

Algorithm 2.7 α update under the energy field gradient u

1. Initialize $\tau = 0$ and k the number of points approximating the path;
2. Compute a new point of the path $\alpha(\frac{\tau}{k}) = \chi_1\left(\alpha(\frac{\tau}{k}); -u(\frac{\tau}{k})\right)$;
3. Project $\alpha(\frac{\tau}{k})$ to $\mathcal{C}$ using algorithm 2.3;
4. Increment $\tau = \tau + 1$;
5. Stop if $\tau > k$; if not, go to step 2.

The last mechanism necessary for geodesic construction is the initialization of a path connecting two elements q_0 and q_1 in $\mathcal{C}$, which is detailed below in algorithm 2.8.

Finally, algorithm 2.9 outlines the computation of a geodesic between two elements $q_0, q_1 \in \mathcal{C}$ using the methods described above for path initialization and straightening.

The length of the geodesic α is then given by:

$$l_\alpha = \int_0^1 dt \ \sqrt{\langle \dot{\alpha}(t), \dot{\alpha}(t) \rangle} \quad . \tag{2.19}$$

Algorithm 2.8 Geodesic initialization

1. Let $\alpha(0) = q_0$, k the discretisation step number along the path, and $\tau = 1$;
2. Compute a unit direction vector field on $\mathbb{S}^2$ between q_0 and q_1: $f = q_1 - \langle q_1, q_0 \rangle q_0$, $f = \frac{f}{\langle f, f \rangle}$;
3. Compute a new element of the path at time $\frac{\tau}{k}$: $\alpha(\frac{\tau}{k}) = \chi_{(\frac{\tau}{k})}(q_0, f)$;
4. Project $\alpha(\frac{\tau}{k})$ to $\mathcal{C}$ using algorithm 2.3;
5. $\tau = \tau + 1$; stop if $\tau > k$; if not, go to step 3.

Algorithm 2.9 Computation of geodesic between two elements $q_0, q_1 \in \mathcal{C}$

1. Initialize a geodesic path α, where $\alpha(0) = q_0$ and $\alpha(1) = q_1$, and a small $\epsilon > 0$, as described in algorithm 2.8;
2. Compute the path velocity vector field $\alpha_t = \frac{d\alpha}{dt}$ using algorithm 2.4;
3. Compute the covariant integral w of α_t as defined in algorithm 2.6;
4. Computer backward parallel transport $\tilde{w}$ of $w(1)$ along the path using algorithm 2.5 (transport a tangent vector w backward, *i.e.* from $\alpha(\frac{\tau}{k})$ to $\alpha(\frac{\tau-1}{k})$);
5. Compute a vector field u giving the gradient field of $E[\alpha]$ and project it to the space $\mathcal{F}_0$ using the expression given by (2.18);
6. Update α in the direction u using the procedure defined in algorithm 2.7;
7. Compute the energy of the current path $E[\alpha] = \frac{1}{2k} \sum_{\tau=0}^{k} \left\langle \alpha_t(\frac{\tau}{k}), \alpha_t(\frac{\tau}{k}) \right\rangle$;
8. The algorithm stops if $\|\nabla E[\alpha]\| < \epsilon$; if not, then go to step 2.

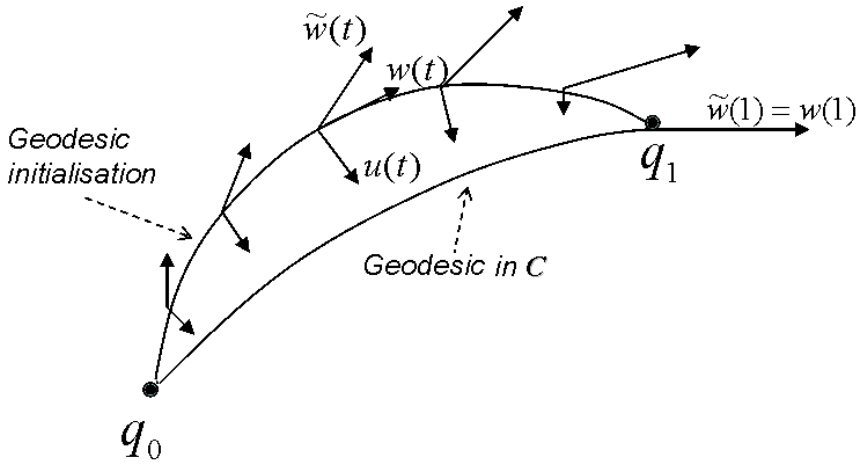

Figure 2.9: Illustration of the path-straightening process.

Geodesic on the elastic shape space

Up to now, we have worked in the space $\mathcal{C}$, called the pre-shape space, where the different elements, in theory an infinite number of elements, represent the same shape. This is due to the fact that the same curve with different starting points, rotated or re-parametrized, is represented by different elements in $\mathcal{C}$. From now on, we will study curves using their representation in the space $\mathcal{C}$ modulo the above mentioned shape preserving curve transformations:

- A shift of the starting point of a curve results in a group action of the unit circle $\mathbb{S}^1$ on $q \in \mathcal{C}$: $p \cdot q(t) = q(t-p) \mod 2\pi, p \in \mathbb{S}^1$. This group acting on an element $q \in \mathcal{C}$ creates an orbit in $\mathcal{C}$. The shortest geodesic path between two elements $q_0, q_1 \in \mathcal{C}$ is thus given by the shortest path between their orbits, which means finding the minimum of the distance $d(q_0, p \cdot q_1)$, which is given by the optimal starting point $\hat{p} = \arg\min_{p\in[0,2\pi]} \langle q_0, p \cdot q_1 \rangle$;

- A rigid rotation of curve is given by a rotation matrix $O = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \theta \in \mathbb{S}^1$ which results in an action of the group $SO(2)$ on an element $q \in \mathcal{C}$ defined as: $O \cdot q(t) = Oq(t), \forall t \in [0, 2\pi)$. The action of $SO(2)$ like that of $\mathbb{S}^1$, creates an orbit in $\mathcal{C}$. Thus, the best match of the elements $q_0, q_1 \in \mathcal{C}$ in the sense of shortest path is given by a rotation matrix $\hat{O} = \arg\min_{O\in SO(2)} \|q_0 - O \cdot q_1\|^2$;

- An arbitrary change of speed along a curve is a curve re-parametrization by a non-linear differentiable map $\epsilon : \mathbb{S}^1 \to \mathbb{S}^1$. Let $\mathcal{E}$ be the space of such elements that preserve the placement of the starting point and the curve orientation, $\mathcal{E} = \left(\epsilon : \mathbb{S}^1 \to \mathbb{S}^1 | \, \dot{\epsilon} > 0\right)$. The re-parametrization results in a group action of $Diff_0^+(\mathbb{S}^1)$ on the element $q \in \mathcal{C}$. This action is defined as $q \cdot \epsilon = \sqrt{\dot{\epsilon}}(q \circ \epsilon)$ and is derived as follows. Let γ_1 and γ_2

be two curves, where $\gamma_1 = \gamma_2(\epsilon)$ and γ_2 is represented by $q \in \mathcal{C}$. We can then write $\dot{\gamma}_1 = \dot{\epsilon}\dot{\gamma}_2(\epsilon) = \dot{\epsilon}q(\epsilon)\|q(\epsilon)\| = \left\| \sqrt{\dot{\epsilon}}q(\epsilon) \right\| \sqrt{\dot{\epsilon}}q(\epsilon)$. Thus, re-parametrizations of q are given by a right action of the group $\mathcal{E}$ on $\mathcal{C}$. The best match between q_0 and q_1 is then given by $\hat{\epsilon} = \arg\min_{\epsilon \in \mathcal{E}} d(q_0, q_1 \cdot \epsilon)$. The method for finding $\hat{\epsilon}$ will be detailed later.

The elastic shape space under the invariance conditions described above, is the quotient space $\mathcal{S} = \mathcal{C}/(\mathbb{S}^1 \times SO(2) \times \mathcal{E})$. The length of the shortest geodesic path between q_0 and q_1 on $\mathcal{S}$ is then given by the minimum distance:

$$d_{\mathcal{S}}(q_0, q_1) = \min_{p \in \mathbb{S}^1, O \in SO(2), \epsilon \in \mathcal{E}} d(q_0, (p \cdot O \cdot q_1) \cdot \epsilon) \tag{2.20}$$

The optimal rotation and starting point giving the minimal geodesic in $\mathcal{C}/(\mathbb{S}^1 \times SO(2))$ are given by the path orthogonal to the orbits of $\mathbb{S}^1$ and $SO(2)$, *i.e.* the tangent vector $\alpha(1)$ is orthogonal to the orbits. The same idea is used to find the shortest geodesic in $\mathcal{S}$, *i.e.* the path orthogonal to the orbits of $\mathbb{S}^1$, $SO(2)$ and of $\mathcal{E}$.

In detail, the method for finding the shortest path between q_0 and q_1 in $\mathcal{S}$ is as follows. Let $\mathcal{E}_{q_1}$ be the orbit of the group $\mathcal{E}$ acting on q_1. The best elastic alignment of two elements is then given by an $\hat{\epsilon}$ obtained when the projection of $\alpha(1)$ on the tangent space $T_{q_1}(\mathcal{E}_{q_1})$ is equal to zero. To solve this optimization problem, Joshi *et al.* [Joshi *et al.*, 2007b] propose to apply a gradient descent technique in the tangent space $T_{q_1}(\mathcal{E}_{q_1})$. More exactly, due to the unknown nature of $T_{q_1}(\mathcal{E}_{q_1})$, the gradient descent is performed in the tangent space to $\mathcal{E}$ at the identity $id = s$, $T_{id}(\mathcal{E}) \subset \mathbb{L}^2(\mathbb{S}^1)$. Then, let $\psi_t : T_{id}(\mathcal{E}) \to \mathcal{E}$ be a flow at the identity in $\mathcal{E}$ along a tangent vector g with $\psi_0(id, g) = s$ for any tangent vector $g \in T_{id}(\mathcal{E})$. This is used to perform gradient descent in a given tangent direction g in the tangent space at the identity $T_{id}(\mathcal{E})$ and to obtain, thus, a re-parametrization for $q_1(s)$. Now, to construct the tangent space $T_{q_1}(\mathcal{E}_{q_1})$, a group action $\phi : \mathcal{C} \times \mathcal{E} \to \mathcal{C}$ is defined, where $\phi_\epsilon(q) = \sqrt{\dot{\epsilon}}q(\epsilon)$, and the differential of which $\phi_* : T(\mathcal{C} \times \mathcal{E}) \to T(\mathcal{C})$ maps the tangent vector g to $\phi_*(g) \in T_q(\mathcal{E}_q)$, where

$$\phi_*(g) = q'(s)g(s) + \frac{1}{2}q(s)g'(s), \ s \in [0, 2\pi) \ . \tag{2.21}$$

Define $V = \{v_i\}, i = 1 \dots m$, the Fourier basis of $T_{id}(\mathcal{E}) \subset \mathbb{L}^2(\mathbb{S}^1)$. The basis in the tangent space $T_{q_1}(\mathcal{E}_{q_1})$ can thus be constructed using $\phi_*(V)$. The projection of the tangent vector to the path $\alpha(1)$ on the tangent space $T_q(\mathcal{E}_q)$ is given by:

$$\pi(\alpha(1)) = \sum_{i=1}^{m} \langle \alpha(1), \phi_*(v_i) \rangle \, \phi_*(v_i) \ . \tag{2.22}$$

Then, $g \in T_{id}(\mathcal{E})$ is constructed using ϕ_*^{-1}, and the flow ψ_t is computed in this direction and gives a re-parametrization ϵ. The process is iterated while $\langle \pi(\alpha_t(1)), \pi(\alpha_t(1)) \rangle > 0$.

Algorithm 2.10 summarizes the computation of a geodesic between $q_0, q_1 \in \mathcal{S}$.

Algorithm 2.10 Main algorithm: computation of a geodesic between $q_0, q_1 \in \mathcal{S}$

1. Initialize $\delta > 0$;
2. Compute the geodesic $\alpha(t)$ between $q_0, q_1 \in \mathcal{C}/(\mathbb{S}^1 \times SO(2))$;
3. Construct the Fourier basis $V = \{v_i\}, i = 1 \dots m$ of $T_{id}(\mathcal{E})$;
4. Project $\alpha(1)$ to $T_{q_1}(\mathcal{E}_{q_1})$ using equation (2.22);
5. Compute $\langle \pi(\alpha_t(1)), \pi(\alpha_t(1)) \rangle$; if it is smaller than δ, then stop; if not, then go to the next step;
6. Construct the tangent vector $g = \sum_{i=1}^{m} \langle \alpha(1), \phi_*(v_i) \rangle v_i$;
7. Compute $\tilde{\epsilon} = \psi_\delta(id, g) = id - \delta g$;
8. Compute the new $q_1 = q_1 \cdot \tilde{\epsilon} = \sqrt{\tilde{\epsilon}'} q_1 \circ \tilde{\epsilon}$;
9. Go to step 2.

Experimental results

Figure 2.10 shows examples of paths with their associated distances for tree crown contours of different species. These experiments were done for tree crowns that chosen to be very characteristic of the species shapes. We can see that the geodesic distance for the same species (the top geodesic between two spruce trees) is a little smaller than for trees of two different species (the geodesics in the middle and the geodesic at the bottom of the figure).

But, as already mentioned, tree crown shapes vary a lot within a class, which can be due, for example, to the age of the tree: a young pine tree has a spiky shape similar to that of spruce, or an old spruce tree can have a crown with dense branches, which results in a quite circular shape similar to pine crown shapes. Let us look now at figure 2.11, which shows geodesics and their distances for trees of different species but with similar shapes. We can see that their shape similarity is reflected in the geodesic distance.

We conclude that shape alone is not sufficiently discriminative for tree classification from 3cm resolution images, due to shape complexity and variety within a species. However, it could be helpful in combination with radiometry and texture based information.

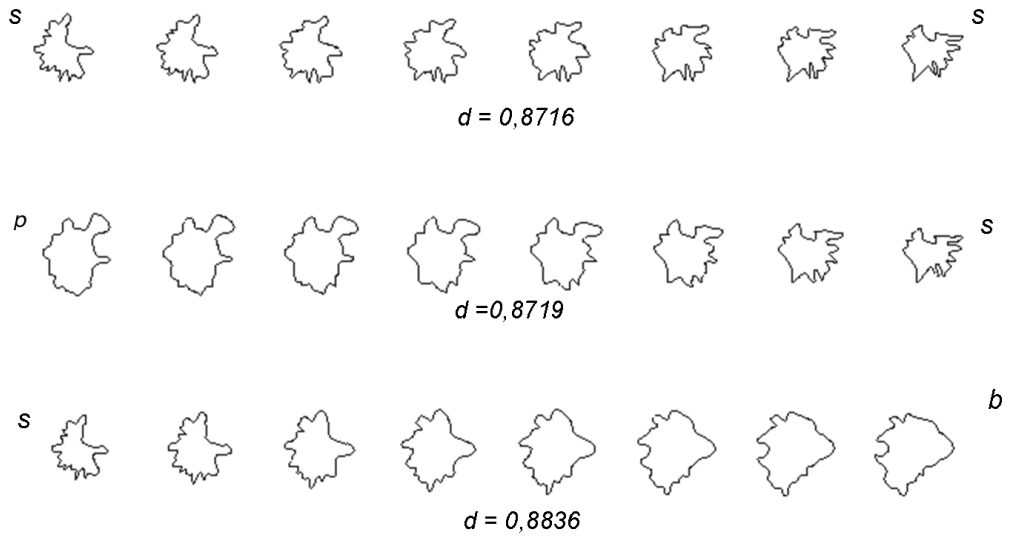

Figure 2.10: Examples of geodesic paths for trees ((s) spruce, (b) birch, (p) pine) using the elastic q-function representation.

2.2 Two metrics: comparison and interesting facts

Here, we present some experimental results comparing the two metrics presented in this chapter. Let us look first at figure 2.12, which shows geodesics between tree crown shapes using the θ angle-function representation versus the elastic q-function representation.

The analysis of the intermediate shapes along the geodesics using θ shows that the transformation passes through a shape more regular and circular, being thus more similar to that of birch. This explains why the birch tree is the species closest to the individuals (cf section 2.1.1). This is due to the bending nature of the metric, meaning that it is less costly to first unfold (to some degree) the curve and then to bend it again. [1]

Looking now at the geodesics under the elastic representation for the same species, we can see that the paths no longer pass through roughly circular curves, because the elastic metric allows stretching as well as bending.

To finish, we present some more interesting experimental results. Figure 2.14 shows geodesic examples comparing the two metrics applied to rectangularly shaped curves, which could, for example, represent building contours in remote sensing images.

[1] This means that more regularly shaped species like birch or pine should be closer to a circle shape than aspen or spruce tree crown shapes. Figure 2.13 shows the geodesic paths and the distances between a spiky spruce tree shape and a circle and a more smoothly shaped birch and a circle.

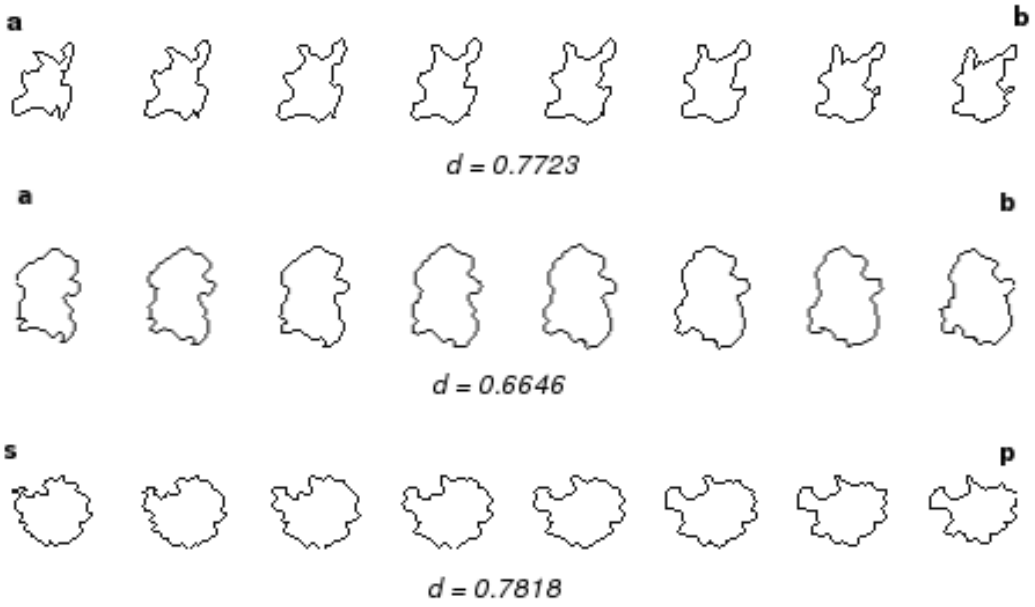

Figure 2.11: Examples of geodesic paths for trees of different species ((a) aspen, (s) spruce, (b) birch, (p) pine) using the elastic q-function representation.

These experiments show that the elastic q-representation, in contrast to the angle-function representation, preserves the straight lines and corners. If we look at the intermediate shapes, we can see that they are more natural in the sense that like the initial and final curves they remain rectangular. From this point of view, we made another set of tests to verify if these properties are preserved for a number of curves within a class. The experiments were done for the same class (of building shaped objects), *i.e.* the objects with boundaries described by rectangular curves in the form of: "L", "U" and "I" are considered to be one class. An intermediate shape for a number of sample shapes is represented by their mean shape in the appropriate shape space. As we know, the geometry of the space of elastic shapes is non-linear, but at every point the tangent space can be constructed, giving the possibility to define statistics in the tangent linear space at the given point. To compute the mean shape Klassen *et al.* [Klassen *et al.*, 2004] suggested the following approach. Consider a set of samples $\{\gamma_1, \gamma_2, \ldots, \gamma_n\}$ and their corresponding q-representations $\{q_1, q_2, \ldots, q_n\}$. The mean element q_μ, and thus the mean curve μ, is given by $q_\mu = \arg\min_q \frac{1}{2n} \sum_{i=1}^{n} d(q, q_i)^2$, and can be found using gradient descent. We are given an energy $E = \delta \frac{1}{2n} \sum_{i=1}^{n} d(q_\mu, q_i)^2 = \frac{\delta}{2n} \sum_{i=1}^{n} \left\langle g_\mu, g_i \right\rangle^2$, and some initial q_μ, where small

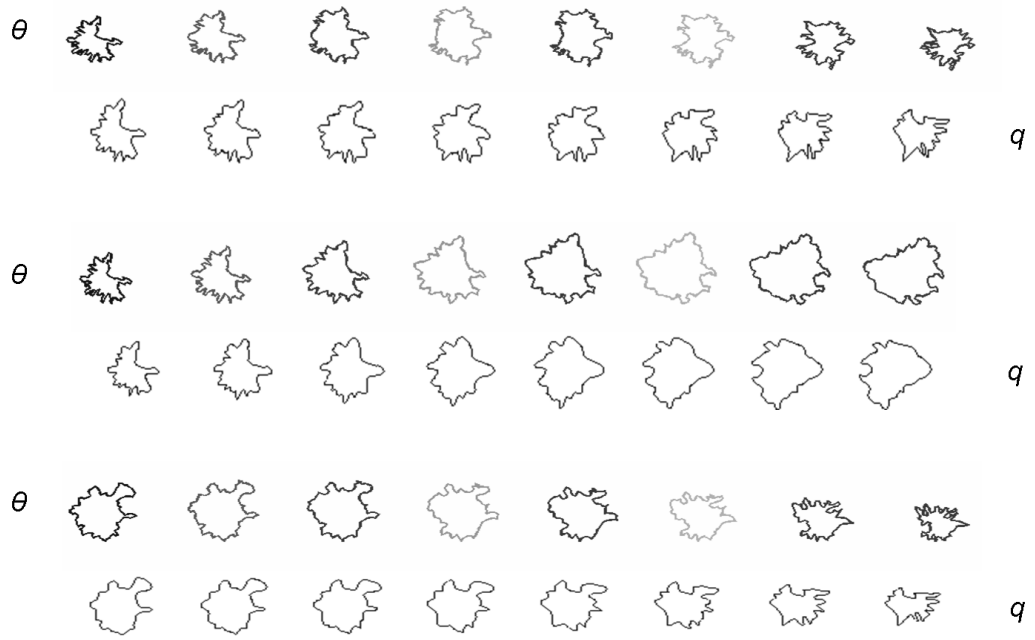

Figure 2.12: Examples of geodesic paths between tree crown shapes using the θ angle-function representation versus elastic the q-function representation.

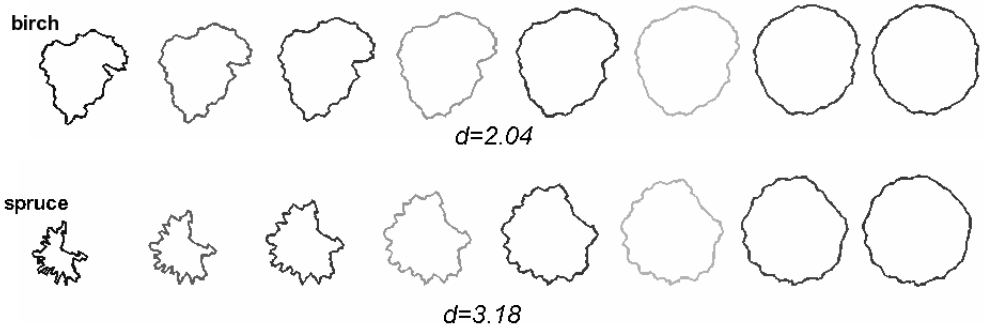

Figure 2.13: Transformation of tree crown curves into a circle along a geodesic path using the bending metric.

$\delta > 0$ and $g_i \in T_{q_\mu}(\mathcal{A})$ are the tangent vectors to the shortest geodesics between q_μ and the elements q_i. The gradient of the energy is then given by $\nabla_{q_\mu} E = \delta \bar{g} = \frac{\delta}{n} \sum_{i=1}^{n} g_i$. The next element can be then computed using the flow $\psi_{0.5\delta}(q_\mu, \bar{g})$. The algorithm stops when $\|\bar{g}\|$ is smaller then some small value giving, thus, a corresponding mean shape (cf Fig. 2.15).

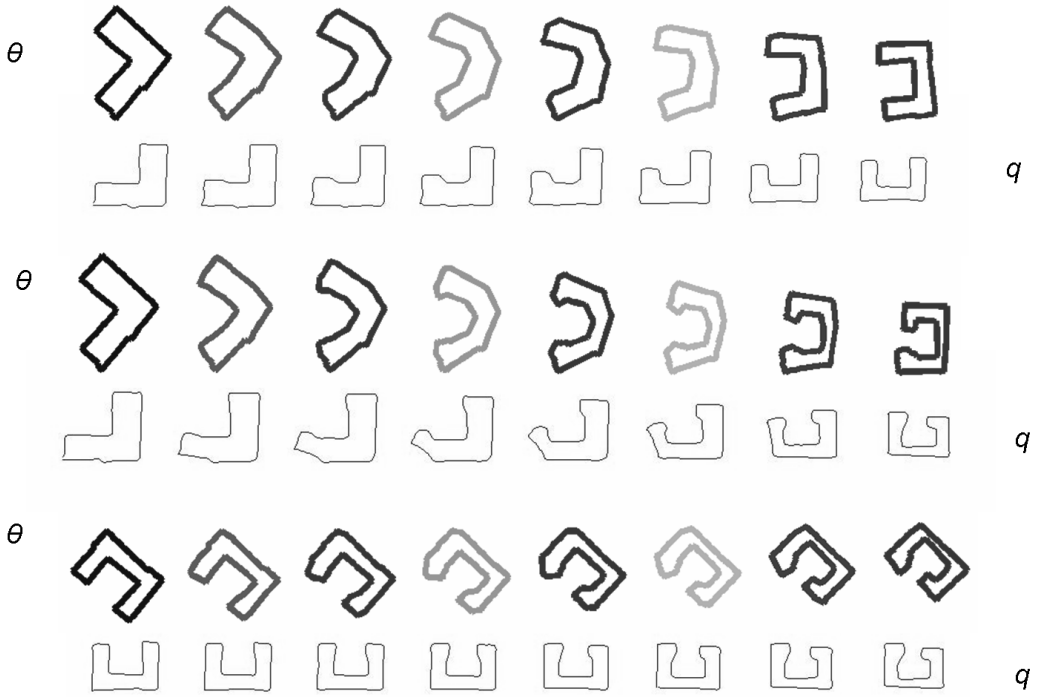

Figure 2.14: Examples of geodesic paths between rectangular shaped curves using the θ angle-function representation versus the elastic q-function representation.

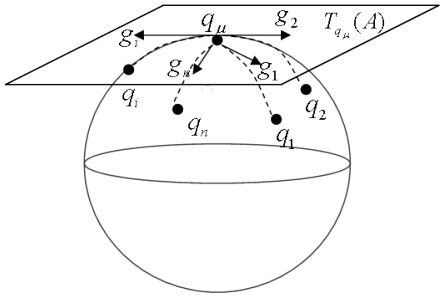

Figure 2.15: Tangent hyper-plane to the sphere at the point q_μ representing the mean shape μ of the samples $\gamma_1 \ldots \gamma_n$.

Figure 2.16 shows examples of mean shapes for building shaped objects, showing that for a whole set of objects of "L", "U" and "I" shapes defined as one class, the mean shape preserves the straight line boundaries and sharp corners.

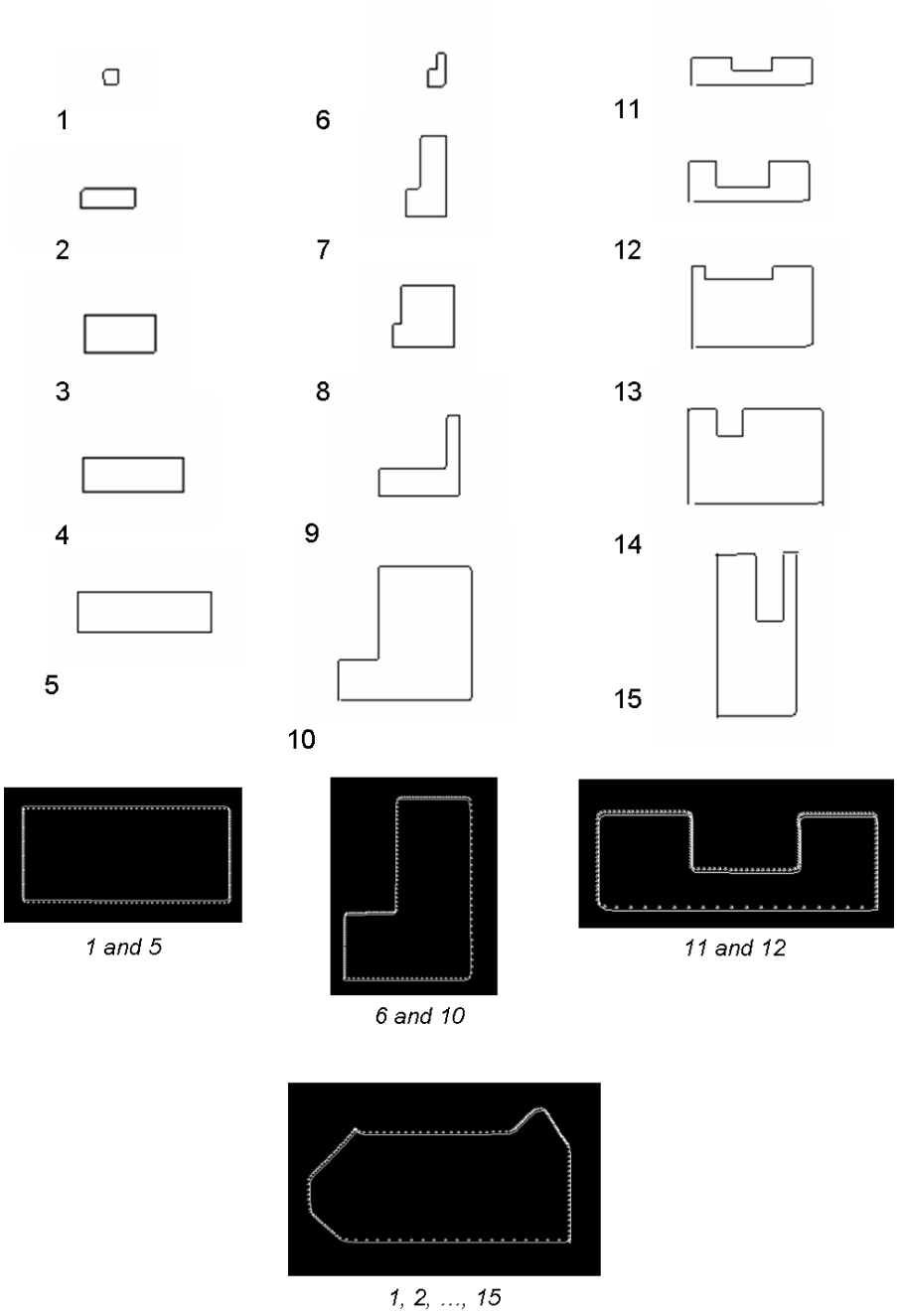

Figure 2.16: 15 ‘building’ shape objects and examples of mean shapes (black background).

These results show that the elastic metric is more suitable for shape analysis because the curve deformation along geodesics under this representation between two similarly shaped curves passes through shapes close to those at the geodesic extremities. In other words: for curves from the same class, in the sense of shape similarity, the geodesic passes through curves from this class.

Chapter 3

Tree species classification using radiometry, texture and shape based features

Descriptors carrying information on the species about the trees combined in training algorithm would allow the classification of individuals.

As mentioned in Introduction, 48 tree crown contours were selected for our study. The contours were carefully delineated manually to preserve important tree crown shape information. SVM classification was performed using a Gaussian kernel. Since the database is quite small, for each experimental run, 50% of the samples were picked at random to form the training set.

First, classification was performed using only radiometric features. Then, texture features were added and finally, shape descriptors were included. The performance was calculated after each step. To evaluate the performance, the average performance, P, of a set of experiments was computed. 5% of the values at the high and low end of the performance scale were excluded from this calculation. P can be expressed as follows

$$P = \left(\sum_{i=1}^{N_e} P_i - \sum_{w=1}^{N_w} P_w - \sum_{b=1}^{N_b} P_b \right) / (N_e - N_w - N_b).$$

$P_i = N_c^i/N$ is the performance of experiment i, N_e the number of experiments, N_c^i the number of correctly classified trees, N the number of trees in experiment i, N_w and N_b the numbers and P_w and P_b the experiment performances of the worst and the best results respectively. The maximum performance P_{max} is also given (from which the best combination for training and test sets can be obtained). Then, we present the *confusion matrix* for the best performance run. A confusion matrix is a visual representation of actual versus predicted classifications.

3.1 Radiometry based features

For vegetation and land use monitoring, colour infra-red film offers a richer set of information than natural colour film. The false colours that describe the three channels are Red for the NIR, Green for the Red, and Blue for the Green channels, [Perrin, 2006]. CIR photography is commonly used to discriminate live healthy trees from dying vegetation.

In our study, CIR allows us to distinguish conifers from deciduous trees. As pointed out in [Erikson, 2004a], the 4 classes of trees - aspen, birch, spruce and pine - can be easily identified as deciduous or coniferous from first order statistics (the mean and standard deviation, computed from the histogram of pixel intensities on the image). This is due to the fact that deciduous trees reflect a substantially greater percentage of infra-red light, cf figure 3.1.

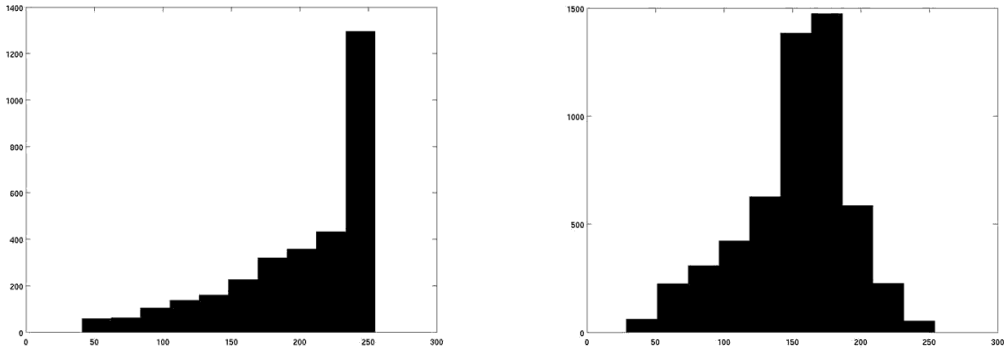

Figure 3.1: Infra-red profiles. Left: a deciduous tree (aspen). Right: a coniferous tree (pine).

Classification based exclusively on these descriptors gives low performance with the average performance being $P = 0.54$ and the maximum, $P_{max} = 0.67$. The confusion matrix for one of the best values of P after repeated classification runs is shown below

$$\begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.167 & 0.666 & 0.167 & 0 \\ 0 & 0.334 & 0.666 & 0 \\ 0 & 0 & 0.167 & 0.833 \end{bmatrix} \begin{matrix} a \\ b \\ c \\ d \end{matrix}$$

where (a) refers to aspen, (b) to birch, (c) to spruce, and (d) to pine. This matrix shows that

there are a substantial number of false positives. For example, 50% of aspens are classified as birch, while 33% of the birch trees are classified as either aspen or spruce. This classifier is especially weak in differentiating within coniferous and deciduous classes, *i.e.* between aspen and birch or between spruce and pine.

3.2 Texture based features

To further distinguish within the deciduous and coniferous classes, texture analysis using grey level co-occurrence matrices (GLCM) is performed. A co-occurrence matrix is a two-dimensional quantitative representation of spatial relationship [Haralick, 1978], well adapted for characterising micro-textures.

Let $\{I(x, y),\ 0 \leq x \leq N-1,\ 0 \leq y \leq N-1\}$ denote an $N \times N$ image with G grey levels as described in [Chen *et al.*, 1998]. The $G \times G$ grey level co-occurrence matrix $P_{\mathbf{d}}$ is defined as

$$P_{\mathbf{d}}(i, j) = |\{(r, s), (r + dx, s + dy) : I(r, s) = i, I(r + dx, s + dy) = j\}|\,.$$

The entry (i, j) of matrix $P_{\mathbf{d}}$ is the number of occurrences of the pair of grey levels i and j which are a displacement $\mathbf{d} = (dx, dy)$ apart. $|\,.\,|$ is the cardinality of a set. GLCMs are a compact representation of pairs of pixel values in relation to each other. They are an example of the second order statistics as defined by Julesz and several useful features can be computed from them.

Nine such matrices were generated for each tree, each matrix representing a different direction or distance. Two texture features, *energy* and *contrast*, were extracted from the GLCMs:

- *The energy term*

  $$\sum_i \sum_j P_{\mathbf{d}}^2(i, j)$$

  is a measure of the homogeneity of the texture. If the grey level transitions are roughly uniformly distributed, which is the case for birch trees, the energy will be small. Conversely, textures which have dominant grey level, *e.g.* pine, aspen or spruce, transition modes have larger energy values;

- *The contrast feature*

  $$\sum_i \sum_j (i - j)^2 P_{\mathbf{d}}(i, j)$$

  is a measure of the local variation present in an image. We selected this feature, related to the auto correlation, to exploit the distinctive features present in the four types of crown surface. Spruce trees, for example, display a radial pattern while aspen have random light and dark regions.

Since these two features are independent of the size and shape of the crown surface, they represent pure texture characteristics.

The evaluation of classifier performance allows us to determine the optimal pair of parameters (**d** and direction). In practice we chose **d** = 1 and 135 degree direction. By incorporating these two texture features into the classifier, we were able to separate the trees into 4 classes with an average performance $P = 0.71$ and a maximum performance $P_{max} = 0.833$. The confusion matrix for one the best values of P after repeated classification runs is shown below:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.833 & 0 & 0.167 \\ 0.334 & 0 & 0.666 & 0 \\ 0 & 0 & 0.167 & 0.833 \end{bmatrix} \begin{matrix} a \\ b \\ c \\ d \end{matrix}$$

Texture information thus improves the classification results for deciduous trees. This is due the fact that one deciduous species (birch (b)) has a roughly uniform texture in comparison to the other (aspen (a)).

3.3 Shape based features

Some tree crown shape analysis was first done to determine the information that might allow us to classify the species. Figure 3.2 shows that aspen crown contours, for example, have an irregular structure; we can also see, the convexities/branches sticking out of the body of the crowns. The branches of spruces are more regular and are oriented more radially. The birch and pine crown contours are more circular, birches being the smoothest. [1]

3.3.1 Tree crown shape representation

To further improve the previous classification results, we propose to study tree crown shapes using the representation of planar shapes by their angle functions, cf [Klassen *et al.*, 2004].

The function θ seems well adapted to exhibit the tree crown shape differences described at the beginning of this section. This can be seen especially in the graphs in $\tilde{\theta} = \theta - \theta_0$ (cf figure 3.2): by substracting from the angle function the line $\theta_0(s) = s$, which represents a unit circle, one obtains an element of the space $\mathbb{L}^2$ which reflects the shape of the curve.

[1] Nevertheless, despite the differences outlined earlier, a significant number of contours remain unclassifiable by human beings just looking at the tree crown shapes, because of variability within a single class.

We, thus propose to define the descriptors based on the functions θ and $\tilde{\theta}$.

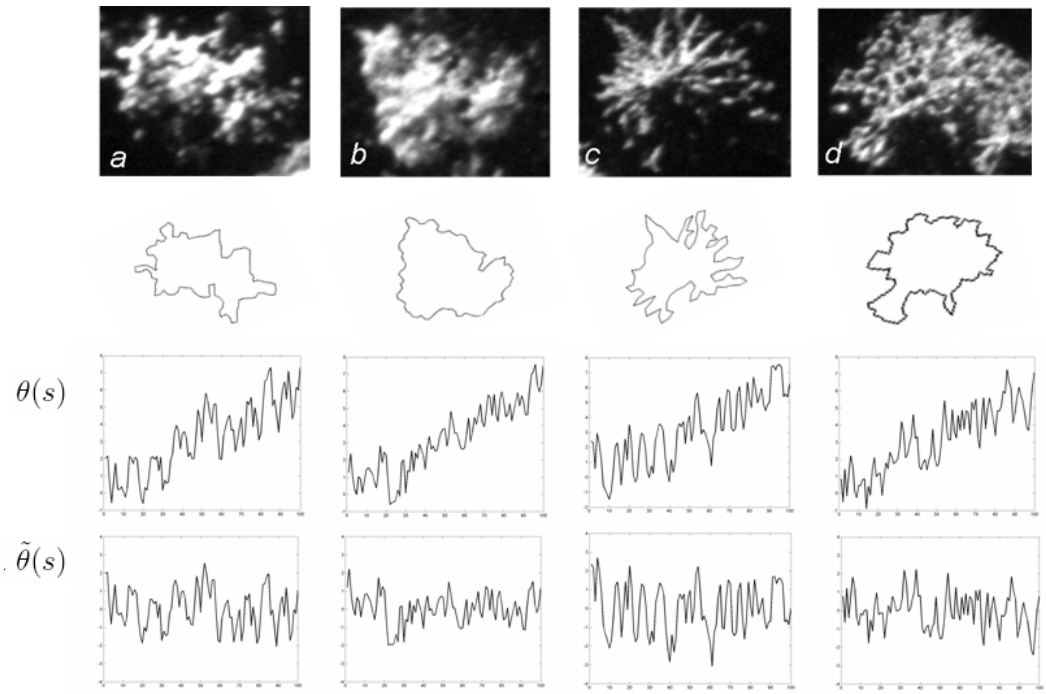

Figure 3.2: Examples of crowns of four species, their contours, the corresponding angle functions θ, and $\tilde{\theta} = \theta - \theta_0$ at the bottom. (a)Aspen, (b)Birch, (c)Spruce, (d)Pine.

We will now describe the tree features created using this representation of the crown contours representing the information necessary for classifying the species.

3.3.2 Tree crown shape appearance

- *Geodesic distance to a circle*

  Using our observations of tree crown shapes in the angle-function representation as described in chapter 2, we note in particular the fact that the geodesics always passe through shapes more circular than the initial and the final curves; and that, pine and birch tree crown shapes are closer to circle, than those of aspen or spruce trees. The first feature is a geodesic distance to a circle $d(\theta, \theta_0)$, which is computed using the geodesic on shape space under the angle-function representation, cf chapter 2, section 2.1.1. For the given shapes $\theta_i, i = 1, ..., M$, we create a vector of "distances to a circle":

$$v_c(\theta) = \{d(\theta_i, \theta_0), \quad i = 1, 2, \ldots, M\} \in \mathbb{R}\,,$$

where M is the number of trees.

We also include geometrical descriptors based on $\tilde{\theta} = \theta - \theta_0$, as follows.

- *Elasticity of crowns*

  Firstly, we translate the property that the crowns of some species have a more regular structure than others, by a measure of contour elasticity:

$$v_e(\tilde{\theta}_i) = \int_0^{2\pi} \dot{\tilde{\theta}}_i(s)^2 ds,$$

$$v_e(\tilde{\theta}) = \{v_e(\tilde{\theta}_i), \;\; i = 1, 2, \ldots, M\} \in \mathbb{R} \, .$$

- *Number of convexities (branches, leaves) of crowns*
  Spruces generally have branches/convexities that are fewer in number and larger than the convexities of birch crowns. This criterion is reflected in number N of local maxima of the angle function.

$$v_N(\tilde{\theta}_i) = N_i,$$

$$v_N(\tilde{\theta}) = \{v_N(\tilde{\theta}_i), \;\; i = 1, 2, \ldots, M\} \in \mathbb{R} \, .$$

- *Size of crown contour irregularities*
  Pine crown shapes are quite close to those of birch, but certain irregularities are larger, with some branches sticking out. We can quantify crown contour irregularities due to the branches, leaves and shadows:

$$v_\mu(\tilde{\theta}_i) = \bar{\mu}_i = \frac{1}{2\pi} \int_0^{2\pi} \left| \tilde{\theta}_i(s) \right| \, ds,$$

$$v_\mu(\tilde{\theta}) = \{\bar{\mu}_i, \;\; i = 1, 2, \ldots, M\} \in \mathbb{R} \, ,$$

$$v_{Var}(\tilde{\theta}_i) = Var_i = \frac{1}{2\pi} \int_0^{2\pi} (\bar{\mu}_i - \left| \tilde{\theta}_i(s) \right|)^2 \, ds,$$

$$v_{Var}(\tilde{\theta}) = \{Var_i, \;\; i = 1, 2, \ldots, M\} \in \mathbb{R} \, .$$

- *Comparisons of crown contour convexities*
  Finally, we calculate the histograms of $\tilde{\theta}$, $h_{\tilde{\theta}}(x) = \int_0^{2\pi} \delta(x, \tilde{\theta}(s)) \, ds$, $x \in \mathbb{R}$, and then Euclidean distances $d(h_i(x), h_j(x)) = \left\| h_i(x) - h_j(x) \right\|$. Then we create the following features associated to each shape, one distance for every class l. The distance is calculated as the distance to the nearest element of the class l:

$$d^l_{min}(h_i) = \min\{d(h_i, h_j)\}^{j=1,\ldots m_l} \in \mathbb{R},$$

$$v_d(\tilde{\theta}) = \{d^l_{min}(h_i),\ \ i = 1, 2, \ldots, M,\ l = 1, \ldots, N_l\} \in \mathbb{R}\,,$$

where N_l is the number of classes and m_l is the number of shapes of each class in the training set.

Now we form vectors using all the features (based on the radiometry, texture and shape information) and apply SVM on these feature vectors. The average performance was $P = 0.747$ and the maximum performance was $P_{max} = 0.87$. An example of a confusion matrix is:

$$\begin{bmatrix} 0.833 & 0.167 & 0 & 0 \\ 0.167 & 0.833 & 0 & 0 \\ 0 & 0.167 & 0.833 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} a \\ b \\ c \\ d \end{matrix}$$

Thus, shape based features allow us to improve the classification of the species within the conifer and deciduous tree groups.

3.4 Summary

In this chapter, we considered the problem of tree species classification from high resolution aerial images based on radiometry, texture and the shape of tree crowns. The images represent forest zones, where the tree crowns were pictured from almost vertically above. The classification was performed on the four most common forest species in Sweden. For our experiments, 48 crowns were delineated manually in order to preserve their shape details. A set of features were then created using the radiometric, texture and shape characteristics of the tree crowns. Classification was performed using a Support Vector Machine with a Gaussian kernel $K(x, x') = \exp(-\frac{\|x-x'\|^2}{2\sigma^2})$.

We have shown that the performance of a classifier based on conventional spectral and texture characteristics can be improved by including shape descriptors in the feature set. By incorporating shape features, the mean classification performance improved by about 4% for 6 samples per class while the maximum performance was 87.5% (Fig. 3.3). To create the new descriptors, the shapes of crowns were analysed using the angle function representation. The representation preserves the characteristics that associate a tree crown with one class or another (cf the discussion in 3.3). The limits imposed by the database size did not give us much freedom to experiment with the training to test sample size ratio. We did, however, observe that the performance mean tended to increase when samples were added to the training set.

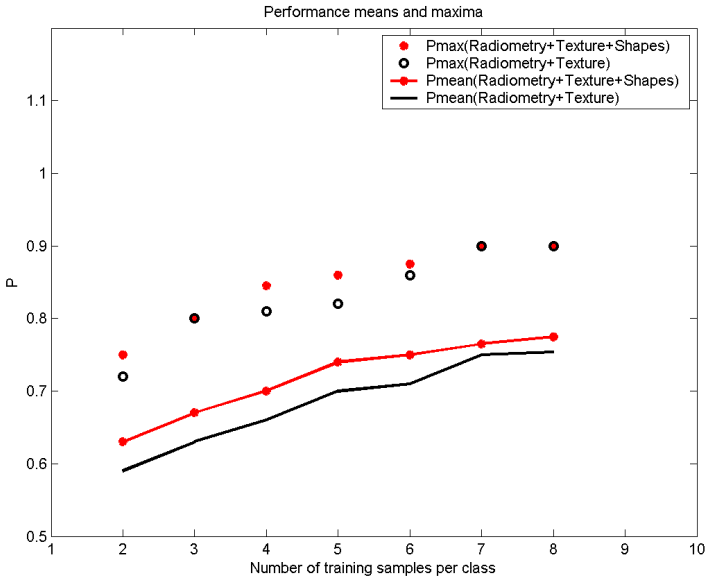

Figure 3.3: Performance means and maxima.

Preliminary to the classification of tree crowns from images, their detection and then segmentation as objects from these images is necessary. At the resolution of the images that we used for study, their geometry is clearly visible (the deciduous trees were identifiable almost to the leaf level), and should be taken into account for associating a tree with a species. Additionally, forest zones represent scenes of high segmentation complexity due to the fact that the trees are in general situated close to each other and may even overlap. This is an important issue: overlapping objects should be segmented as a set of individual objects and not as a one single object. This is an especially challenging task for trees of the same species grouped together (even for manual extraction by an expert). We are thus interested in developing a method that allows authomatic extraction of multiple objects from images of very high resolution, the images representing complex scenes with overlapping objects.

Stochastic marked point processes are a technique for multiple object extraction known for their ability to include geometrical information about the objects sought as well as their interactions in the scene. Marked point process models have been successfully applied to the extraction of objects from images of lower resolution, the objects thus having simple visible geometry [Lacoste *et al.*, 2005, Perrin *et al.*, 2005, Ortner *et al.*, 2007, Descamps *et al.*, 2009]. The second topic of this dissertation, presented in the following part, consists in the generalisation of the marked point point process approach to complex arbitrarily-shaped multiple object extraction from very high resolution images.

Part II

Marked Point Processes using Active Contours and Shape Priors

Chapter 4

Introduction

Object extraction 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 military, *e.g.* detection on the ground of strategically important objects; and civilian, like illegal construction detection [Bayburt *et al.*, 2008]; tree counting [Gougeon, 1995b, Perrin, 2006, Horvath, 2007] and species classification [Leckie *et al.*, 2003, Erikson, 2004a, Kulikova *et al.*, 2007] for biomass or biodiversity estimation; or bird counting for monitoring population changes [Descamps *et al.*, 2009].

The problem of object extraction in image analysis can be formulated as an inverse problem, which consists in obtaining some model parameter value r from the observed data D. Inverse problems are typically solved as optimization problems, where the minimum of an energy function $H(r)$ in the space of model parameters is looked for. The energy in the Bayesian framework is traditionally composed of two terms: a prior or a regularization term $H_p(r)$, which contains the knowledge about the structure or the behavior of r; and a data term or a likelihood term $H_D(r)$, which relates the parameter and the data D. The parameter is then usually estimated using a Maximum A Posteriori (MAP) estimate and is given by $\hat{r} = \arg\min_r \Big(H_p(r) + c_0 H_D(r)\Big)$, with c_0 a weighting parameter.

Nowadays, the resolution of aerial images is approaching a few centimetres (cf figure 4.1). As a result, the geometry of objects in the scene needs to be taken into account for accurate object extraction. Stochastic point processes, in particular marked point processes, are known for their ability to model statistically geometrical objects. Additionally, the point process framework can take into account information about the spatial repartition of objects in the image scene by modeling inter-object interactions. A probability distribution is defined on the configuration space of an arbitrary number of objects, which depends on the relation between the objects and the image, and on the configuration of individual objects and their joint relations. The extracted objects are then those in the optimal configuration. To find the optimal multiple-object configuration, a MAP estimate is usually used.

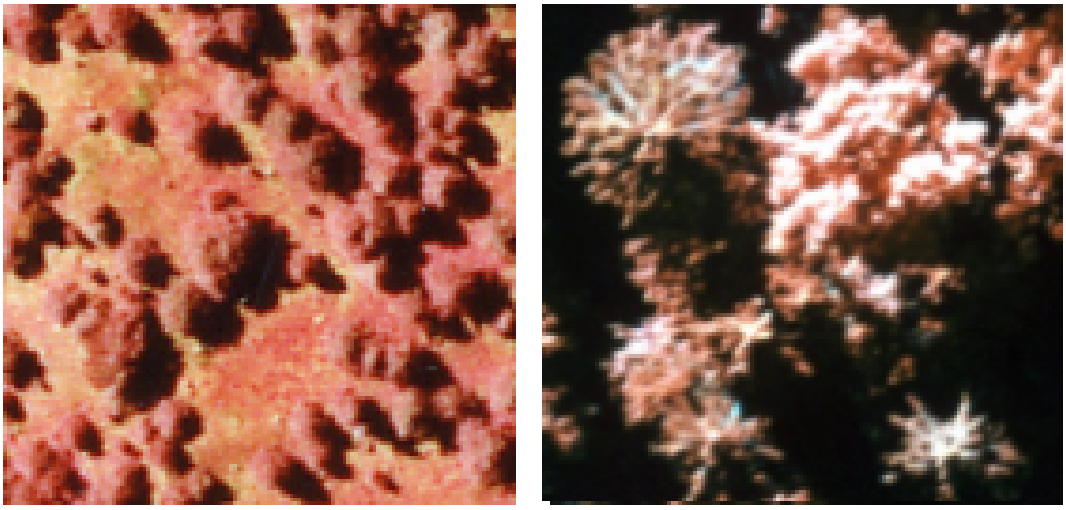

Figure 4.1: Left: fragment of a CIR image of several tens of cm resolution representing the top parts of tree crowns, © IFN. Right: fragment of a CIR image of 3 cm resolution representing the top parts of tree crowns, © CBA, Uppsala.

In previous work, the objects involved have been represented using simple geometrical shapes, *e.g.* discs, ellipses, or rectangles. The resulting models have been applied to the extraction of different types of object from remote sensing images, *e.g.* road networks [Stoica *et al.*, 2000, Lacoste *et al.*, 2005], trees [Perrin *et al.*, 2005], buildings [Ortner *et al.*, 2007], and flamingos [Descamps *et al.*, 2009], but the simplified nature of the individual objects limits the geometrical precision that can be achieved.

The aim of this work is to lift this geometrical restriction. One could define the distribution on the space of all closed curves, but as we see from the previous part, the spaces of such objects are non-linear manifolds with complex topology. Defining a proper metric on these spaces is not straightforward. Therefore, describing even a simple dynamics like diffusion, for one object, becomes a challenging problem. Instead, we propose to reduce the object space to a subset of objects that are locally adapted to the data. Each object is a local minimum of the single-object energy associated with an active contour. We then use the full energy including interactions to find the optimal multiple-object configuration composed of these locally adapted objects.

The single-object space considered is thus still of small dimension, but the possible individual objects are determined not *a priori*, but by the image data and a single-object version of the model. As a result, they can be **arbitrary closed curves**. Once the single-object

space is defined, we define then a Gibbs energy, and hence a probability distribution, on the configuration space of an arbitrary number of objects. We present this model in chapter 6. The model is called "Model I" in appendix B, where the calculation details are given.

The second aim of this work is a further extension of the MPP framework, which consists in defining a method for incorporating strong prior shape information into Model I. The principles remain the same as those of Model I, but the single-object space is defined in such a way as to take into account prior knowledge about the shape of the objects that we wish to extract. This part is detailed in chapter 7. The model is called "Model II".

To find the MAP estimate using the full energy, we sample using a multiple birth-and-death process embedded in an annealing scheme [Descombes *et al.*, 2009].

Chapter 5

Point Processes: review

Point processes are widely used for modeling and analysing spatial data in domains such as epidemiology, ecology and the environment, astronomy and geography, where the data elements (represented by points or objects) have a particular spatial structure, that can also be influenced by object interactions. For example, in environmental applications, like forestry, the trees are close to each other up to some tolerance distance that can be modeled by a repulsion between the objects; or in biology, where this effect is provoked by, for example, the competition for space or food. There is a great deal of work dedicated to this methodology and its applications. In this chapter we recall the main definitions and theorems of point process theory. For more complete information and details, the reader can refer, for example, to the books [Daley and Vere-Jones, 1988, van Lieshout, 2000, Stroock, 2005, Jacobsen, 2006].

5.1 Main definitions

We assume a space χ equipped with a metric d such that (χ, d) is complete and separable. Any element $x \in \chi$ is called a point. A countable, unordered set of points in χ $\mathbf{x} = \{x_1, \ldots, x_n\}$, $n \in \mathbb{N}$ is said to be a configuration (cf figure 5.1).

We consider the configurations $\mathbf{x}$ from a space denoted N^{lf}, such that every $\mathbf{x}$ places a finite number $N_{\mathbf{x}}(A)$ of points in any bounded Borel set $A \subseteq \chi$. These configurations are called locally finite. For the applications of object extraction from images, we consider bounded sets χ. Locally finite configurations are then finite and the associated space is denoted N^f. Let χ be equipped with a Borel measure ν, in general the Lebesgue measure Λ, so that, the product measure ν^n is a measure on χ^n. Then the measurable subsets of N^f containing n unordered points are defined as follows:

$$N_n^f = \left\{ \mathbf{x} \in N^f \ : \ N_{\mathbf{x}}(\chi) = n \right\} \quad .$$

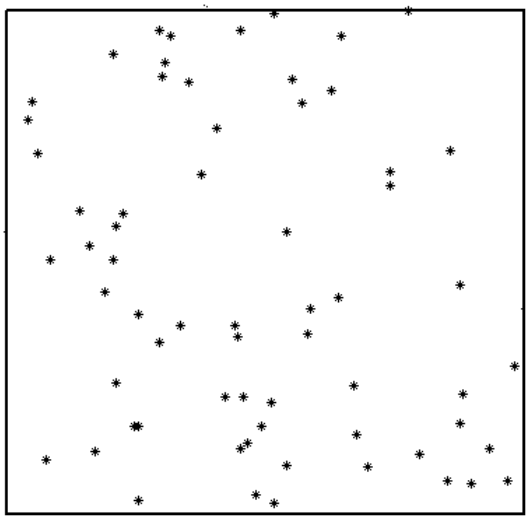

Figure 5.1: A configuration of points on $\chi = [0,1]^2$ with Euclidean metric.

The measure of N_n^f is $\nu(\chi)^n/n!$, the factor $n!$ being needed since χ^n is ordered. Thus, the measure of N^f is given by

$$\nu(N^f) = \sum_{n=0}^{\infty} \frac{\nu(\chi)^n}{n!} = e^{\nu(\chi)} \tag{5.1}$$

A point process, the state of which at a given moment is a random configuration of points is defined as follows.

Definition 5.2. *A **point process** on χ is a mapping X from a probability space $(\mathbf{\Omega}, \mathcal{A}, \mathbb{P})$ into N^{lf} such that for all bounded Borel sets $A \subseteq \chi$, the number $N(A) = N_X(A)$ of points falling in A is a finite random variable.*

When configurations of objects are considered, some random variables describing the object geometry are added to the points. For example, to define a configuration of discs, a radius is added to each point (cf Fig. 5.2). We then have the following definition:

Definition 5.3. *A **marked point process** on $\chi = \mathcal{P} \times \mathcal{M}$ is a point process with positions in $\mathcal{P}$ and marks in $\mathcal{M}$ such that the process of unmarked points is a point process.*

Since χ is bounded, the point process is said to be finite. To obtain a realisation of a point process, we need a discrete probability distribution (p_n), $n \in \mathbb{N}$ for the number of points in the configuration and a family of symmetric probability densities $j_n(x_1, \ldots, x_n)$, $n \in \mathbb{N}$ on χ^n for the positions. The symmetry property is required, since a point process is indifferent with respect to the order of points in the configuration.

The best known example of such a point process is the Poisson process defined as follows.

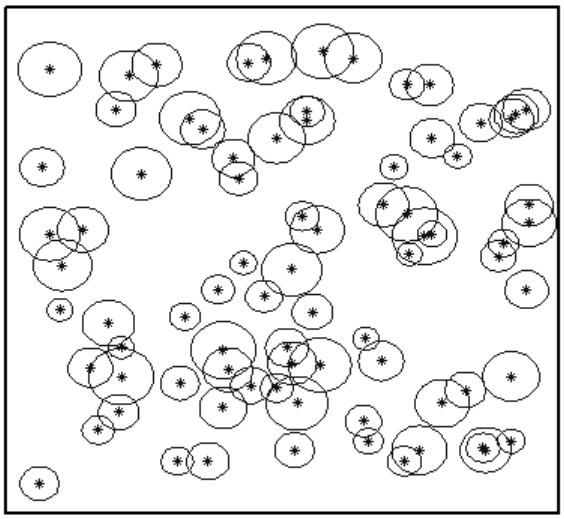

Figure 5.2: A configuration of circles described by their positions on $\mathcal{P} = [0,1]^2$ and radii from $[r_{min}, r_{max}]$ as marks on $\mathcal{M} = [r_{min}, r_{max}]$.

Definition 5.4. *A point process X is a **Poisson process** on χ with intensity measure ν if*

- *N(A) is Poisson distributed with mean $\nu(A)$ for every bounded Borel set $A \subseteq \chi$;*
- *For any n disjoint bounded Borel sets $A_1, \ldots, A_n$ the corresponding random variables $N(A_1), \ldots, N(A_n)$ are independent.*

A Poisson process is called homogeneous if ν is proportional to the Lebesgue measure, such that $\nu(A) = \lambda\Lambda(A)$. In the case of a non-homogeneous Poisson process, a Borel measurable intensity function $\lambda(x)$ on χ is defined as the Radon-Nikodym derivative of ν w.r.t. Lebesgue measure as follows:

$$\nu(A) = \int_A \lambda(x)\Lambda(dx) < \infty.$$

The distribution π_ν of the Poisson process of intensity λ for all $B \in \mathcal{N}^f$ is then written as

$$\pi_\nu(B) = e^{-\nu(\chi)}\left(\mathbf{1}_{[\emptyset\in B]} + \sum_{n=1}^{\infty}\frac{\pi_{\nu_n}(B)}{n!}\right), \tag{5.2}$$

where $\nu_n \equiv \nu^n$, $1_{[\emptyset\in B]}$ is the indicator, which is equal to 1, if $\emptyset \in B$, and equal to 0 otherwise, and

$$\pi_{\nu_n}(B) = \int_\chi \ldots \int_\chi \mathbf{1}_{[\{x_1,\ldots,x_n\}\in B]}\nu(dx_1)\ldots\nu(dx_n). \tag{5.3}$$

In order to construct a particular point process X, we consider its probability density (Radon-Nikodym derivative) w.r.t. the Poisson process which is called the reference process. Let π_ν be the distribution of the Poisson process with intensity measure ν, consider $p : N^f \to [0, \infty)$ a measurable function on the collection of finite point configurations such that

$$\int_{N^f} p(\mathbf{x}) d\pi_\nu(\mathbf{x}) = 1. \tag{5.4}$$

p is then a probability density function and it defines a point process X on χ.

Taking into account (5.1) and (5.2), the distribution of the total number of points of a process X is defined by its density p given as follows

$$p_n = \frac{e^{-\nu(\chi)}}{n!} \int_\chi \cdots \int_\chi p(\{x_1, \ldots, x_n\}) d\nu(x_1) \cdots d\nu(x_n)$$

A class of models that is specially designed to take into account point interactions including the Poisson process are the Markov spatial point processes, known also under the name of Gibbs point processes, since the similar concept of Gibbs point process is used in statistical physics.

Definition 5.5. *Let (χ, d) be a complete, separable metric space, ν a finite non-atomic Borel measure, and π_ν the distribution of the Poisson process on χ with intensity measure ν. Let X be a point process on χ defined by its density p w.r.t. π_ν. Then, X is a **Markov point process** w.r.t. the symmetric, reflexive relation $\sim$ on χ if for every configuration $\mathbf{x} \in N^f$ such that $p(\mathbf{x}) > 0$,*

- *$p(\mathbf{y}) > 0$ for all $\mathbf{y} \subseteq \mathbf{x}$ (heredity);*
- *for all $u \in \chi$, $p(\mathbf{x} \cup \{u\})/p(\mathbf{x})$ depends only on u and its neighbourhood $\partial(\{u\}) \cap \mathbf{x} = \{x \in \mathbf{x} : u \sim x\}$.*

Definition 5.6. *A configuration $\mathbf{x} \in N^f$ is said to be a clique if all elements of $\mathbf{x}$ are neighbours of each other w.r.t. the symmetric and reflexive relation $\sim$. An empty set is also a clique.*

A similar result to the Hammersley-Clifford theorem provides a definition of the density of a Markov point process in terms of interactions between the configuration points.

Theorem 5.2. *A point process density function $p : N^f \to [0, \infty)$ is Markov w.r.t. a neighbourhood relation $\sim$ iff there is a measurable function $\phi : N^f \to [0, \infty)$ such that:*

$$p(\mathbf{x}) = \alpha \prod_{\text{cliques } \mathbf{y} \subseteq \mathbf{x}} \phi(\mathbf{y}) \tag{5.5}$$

for all $\mathbf{x} \in N^f$.

One of the well known Markov point processes is the Strauss process. Its density function w.r.t the reference Poisson point process with distribution π_ν is given as follows:

$$p(\mathbf{x}) = \alpha\beta^{n(\mathbf{x})}\gamma^{s(\mathbf{x})}, \tag{5.6}$$

where $\beta > 0$, $\gamma \geq 0$, $n(\mathbf{x}) = N_X(\chi)$, and $s(\mathbf{x})$ is a number of cliques of order 2 (*i.e.* cliques containing two objects) w.r.t the relation $\sim$ such that $u \sim v \Leftrightarrow d(u, v) < R$. β is the parameter that changes the intensity of the process given as $\beta\lambda$ w.r.t to the reference process; the parameter γ changes the process' behaviour, *i.e.*

- if $\gamma = 1$, the process is a Poisson process of intensity $\beta\lambda$;
- if $\gamma = 0$, two points satisfying $\sim$ are not allowed;
- if $\gamma \in (0, 1)$, there is a repulsion between two points close in the sense of the $\sim$ relation;
- if $\gamma > 1$, the process is a clustered process, *i.e.* there is an effect of point clustering. In this case, the number of objects of the process has to be bounded above, otherwise the density is not integrable and the process is undefined

Another interest of a Gibbs process is that its density function can be written as the exponential of an energy (we give the details in the section that follows). In [Perrin, 2006] several models for forest study are defined. These models are based on Gibbs point processes with interactions between two neighbouring objects, which allow the incorporation of properties such as repulsion or attraction between two objects (two tree crowns in the case of the forest study), depending on, for example, the tree species or the distance between the trees, or a particular structure in the plantation. The energy is defined as the sum of two terms: a term $H_d(\mathbf{x})$ related to the data and a prior term $H_p(\mathbf{x})$, that models the constraints on the objects as well as on their interactions. The interaction constraints are then defined for the cliques of order 2 w.r.t to the given symmetric relations $\sim$. With each of the relations is associated an energy $H_\sim(.)$ (some relations being favoured and others being penalized). The prior energy thus takes the form:

$$H_p(\mathbf{x}) = \sum_i \sum_{u \in \mathbf{x}} H_i(u) + \sum_j \sum_{u \sim_j v;\, u,v \in \mathbf{x}} H_{\sim_j}(u \sim_j v).$$

5.2 Point process simulation

In order to develop a model of object extraction based on a marked point process, a mechanism of sampling from a process is necessary. To introduce the mechanism, we use the notation defined in the previous section. We wish to simulate a Gibbs point process with a distribution π defined by its density function p w.r.t to a reference Poisson process with a distribution π_ν, on the space of configurations of a finite number of elements N^f . The

objects considered are from $\chi = \mathcal{P} \times \mathcal{M}$ with positions in $\mathcal{P}$ and marks in $\mathcal{M}$. Denote the configurations of such objects as $\mathbf{x} = \{u_1, \ldots, u_{n(\mathbf{x})}\}$, where $u_i = [p_i, m_i]$ is an object located at p_i and with mark m_i. For all $\mathbf{x} \in N^f$ and all $B \in \mathcal{N}^f$ the process distribution is given by

$$\pi(B) = \int_B p(\mathbf{x}) \pi_\nu(d\mathbf{x}).$$

The density function of the point process in Gibbs form is written as the exponential of an energy and is given by the expression:

$$p(\mathbf{x}) = \frac{1}{Z} \exp\{-H(\mathbf{x})\},$$

where Z is the normalizing constant

$$Z = \int_{\mathbf{x} \in N^f} \exp\{-H(\mathbf{x})\} \, d\mathbf{x}.$$

The density function includes the normalizing factor Z which is not tractable. Due to this fact, the direct sampling of configurations from the density is impossible. In such a case, an alternative is to use a Markov Chain Monte Carlo (MCMC) type approach, which consists in approximating the process by a Markov chain $X_0, \ldots, X_n, \ldots$ in the space of configurations such that beginning at any configuration, the defined Markov chain converges to the point process. The conditions of Markov chain convergence are listed in the following section; details and additional information about the properties of Markov chains can be found in, for example, [Stroock, 2005] or [Miller and Pankov, 2001] (in Russian).

The idea initially proposed by Metropolis *et al.* , is to run a Markov process converging to a distribution which corresponds to the distribution of the process to be approximated. Nowadays, there exist many different such techniques with diverse kinds of transition operators, mixing degree, and implementation complexity. The most used algorithms are the following ones: a reversible jump version of the Metropolis-Hastings algorithm proposed in [Geyer and Moller, 1994], where birth and death proposed allow the number of objects in the configuration to change; and a generalisation of this algorithm embedding a mixture of several kernels proposed in [Green, 1995], *i.e.* birth and death kernel as well as the kernels which do not change the number of objects in the configuration but accelerate convergence, *e.g.* simple object perturbations (translation, rotation and dilatation); object fusion and division; or birth and death in an object's neighbourhood [Lacoste, 2004, Ortner, 2004, Perrin, 2006]. In these algorithms, the transitions are defined by a new proposed state which is then accepted or rejected w.r.t to its likelihood in comparison to the previous state. The spatial birth-and-death algorithm described in [Preston, 1977, van Lieshout, 1993], avoids the rejection step, but the process stays in a state with an exponentially distributed time, which depends on the total birth and death rates.

For our model, we chose a recently developed **multiple** birth-and-death process [Descombes *et al.*, 2009] for its rapid mixing and because only the death intensity depends on the energy of the current configuration with the birth intensity being constant. We will describe the algorithm in detail in the following chapter.

As we mentioned before, the density of the Gibbs marked point process takes the form of the exponential of an energy $H(\mathbf{x})$. From the application point of view, this energy consists of two terms: a data term, that relates the objects to the data, and a prior term, that introduces prior information through the interaction of objects in the configuration. The data term is negative, when the configuration of objects fits the image well. We look for the best configuration of objects in the sense of minimum of the energy $-H(\mathbf{x})$, which is usually estimated using the maximum a posteriori estimator. In order to do this, the process is embedded in an annealing scheme. The principle is to simulate a non-homogeneous Markov chain (X_n) that converges not to the distribution π, but to the distribution π_β given by $\pi_\beta = p(\mathbf{x})^\beta \pi_\nu(d\mathbf{x})$, where β is a parameter called the inverse temperature of the system, $T = \frac{1}{\beta}$. The temperature T is decreasing slowly during the simulation, and as T tends to 0 the distribution π_β tends to a Dirac distribution on the configuration minimizing the energy.

5.3 Markov chains: reference theorems

In this section we recall the main definitions of Markov chain methodology and the properties of Markov chain convergence to an equilibrium distribution, denoted π, *i.e.* which does not depend on the initial distribution, the necessary condition for independence of initial conditions for MCMC type algorithms.

Definition 5.7. *A **Markov chain** (MC) is a sequence of random variables $S_1, S_2, \ldots, S_k, \ldots$, with the possible values from a countable set $S = \{s_1, s_2, \ldots, s_k, \ldots\}$ called the state space and satisfying the Markov property:*

$$\mathbf{P}(S_{n+1} = s_{n+1} \mid S_n = s_n, \ldots, S_0 = s_0) = \mathbf{P}(S_{n+1} = s_{n+1} \mid S_n = s_n). \tag{5.7}$$

Definition 5.8. *A Markov chain is **(time-)homogeneous** if it does not depend on n,* i.e. *if for all $n > 1$*

$$\mathbf{P}(S_{n+1} = x \mid S_n = y) = \mathbf{P}(S_n = x \mid S_{n-1} = y). \tag{5.8}$$

The probability of going from state s_i to state s_j in one step (single-step probability transition) is defined as follows:

$$p_{i,j} = \mathbf{P}(S_1 = s_j \mid S_0 = s_i), \tag{5.9}$$

and in n step:

$$p_{i,j}^{(n)} = \mathbf{P}(S_n = s_j \mid S_0 = s_i). \tag{5.10}$$

For a homogeneous MC:

$$p_{i,j}^{(n)} = \mathbf{P}(S_{n+k} = s_j \mid S_k = s_i).$$

A Markov chain is characterized by its initial distribution $P(S_0 = s)$. The evolution of the chain in time is then given as follows:

$$\mathbf{P}(S_n) = \sum_m p_{m,j} \mathbf{P}(S_{n-1} = s_m) = \sum_m p_{m,j}^{(n)} \mathbf{P}(S_0 = s_m). \tag{5.11}$$

Consider a time-homogeneous Markov chain. The following properties of a Markov chain guarantee its ergodicity, and the existence of an equilibrium distribution.

Definition 5.9. *The states s_i and s_j are communicating if there exist $m, n > 1$ such that $p_{k,j}^{(m)} > 0$ and $p_{j,k}^{(n)} > 0$. A set of states is called a **communicating class** if every pair of states is communicating.*

Definition 5.10. *An MC is **irreducible** if its state space is a single communication class,* i.e. *any state is accessible from any state.*

Let d_j be a greatest common divisor of the values $\left\{n \geq 1 \ : \ \mathbf{P}(S_n = s_j \mid S_0 = s_j) > 0\right\}$.

Definition 5.11. *A state s_j is called **periodic** with period d_j, if $d_j > 1$; otherwise, it is called **aperiodic**.*

Return to a periodic state is possible only in a number of steps that is a multiple of $d_j > 1$.

Definition 5.12. *An MC is said to be **aperiodic**, if all its states are aperiodic.*

Let $f_j(n) = \mathbf{P}(S_n = s_j, S_{n-1} \neq s_j, \ldots, S_1 \neq s_j \mid S_0 = s_j)$ denotes a probability of a first return to the state s_j in n steps.

Definition 5.13. *A state s_k is said to be **positive recurrent** if the time of return to the state,* i.e. *the random variable T_k with distribution*

$$\mathbf{P}(T_k = n) = f_k(n), \ n = 1, 2, \ldots, \tag{5.12}$$

has a finite expectation

$$M_k = E(T_k) = \sum_{n=1}^{\infty} n f_k(n) < \infty.$$

Theorem 5.3. *If for an MC the conditions of irreducibility and aperiodicity are satisfied, and if the MC has a positive recurrent state, then for any $i, j = 0, 1, \ldots$ the following limits, that do not depend on i, exist:*

$$p_{i,j}^{(n)} \to p_j > 0, as\ n \to \infty. \tag{5.13}$$

The $\left\{p_j\right\}$ comprise the single solution of the following system of equations

$$p_j = \sum_{k=0}^{\infty} p_{k,j} p_k,\ j = 0, 1, \ldots \tag{5.14}$$

$$\sum_{j=0}^{\infty} p_j = 1. \tag{5.15}$$

Definition 5.14. *A Markov chain is **ergodic** if it satisfies the conditions (5.13) - (5.15). The distribution $\pi = \{p_0, p_1, \ldots\}$ is called the **stationary distribution** or invariant measure and it is said to be an **equilibrium distribution**, because it does not depend on the initial distribution.*

A more powerful property of Markov chain convergence is its reversibility.

Definition 5.15. *A Markov chain is called **reversible** if there exists a distribution π such that*

$$\pi_i p_{i,j} = \pi_j p_{j,i}$$

*called the **detailed balance** condition.*

For a reversible MC, π is a stationary distribution since $\sum_i \pi_i p_{i,j} = \pi_j$.

Chapter 6

Marked Point Process for multiple arbitrarily-shaped object extraction

In this chapter we present a marked point process model for multiple arbitrarily-shaped object extraction. We first define a probability distribution for multiple-object configurations, where each individual object is represented by its boundary, i.e. a closed curve in the image domain. The distribution is the Gibbs distribution corresponding to an energy defined on the configuration space of an arbitrary number of objects. This energy consists of a sum of single-object energies, plus an interaction term that penalizes object overlap. The single-object energy is the sum of a term that enforces boundary smoothness, and a data term relating an object to the image. In the absence of an interaction term, the MAP estimate would thus consist of the subset of these single objects with negative energy. This would probably lead to degenerate solution configurations, however, and so an interaction term is added that controls the relation between different objects, in particular discouraging overlaps.

The individual objects are defined by evolving a number of initial curves under gradient descent to local minima of the single-object energy. The initial curves are circles centred on each image pixel, with radii in a certain range, the radius being the mark associated to each point. The size of the single-object space is thus not greater than that of the set of initial curves. This is similar to the object sets used in previous point process models preserving thus computational efficiency.

To find the optimal multiple-object configuration, we compute a MAP estimate by sampling from a multiple birth-and-death process embedded in an annealing scheme. The birth step samples a number of initial circle centres from a uniform Poisson process, with uniformly sampled radii. The circles are added to the current configuration, and then evolved to local minima of the single-object energy, producing a configuration of multiple locally adapted objects. The death step then removes a number of objects with a probability that

depends on the temperature-weighted energy difference between the configurations with and without each object. Iteration is stopped if all the objects added in the birth step and only these are removed in the following death step. The performance of the approach is demonstrated via experimental results on synthetic and real data.

The use of a birth-and-death process allows the number of objects to be unknown a priori. Our approach can be thus thought of as an extension of the active contour methodology [Kass *et al.*, 1998] to an *a priori* unknown number of objects. A great deal of work has been done within the active contour framework using the distance-function level set representation [Caselles *et al.*, 1997, Sethian, 1999, Leventon and Grimson, 2000, Osher and Fedkiw, 2003, Cremers *et al.*, 2006a]. This representation allows arbitrary topology, *i.e.* an arbitrary number of contours, at least in principle. However, it is a representation of a region with arbitrary topology, not an arbitrary number of distinct objects: overlapping objects, for example, cannot be represented. Second, the algorithms used are forms of deterministic gradient descent, meaning that the result may be very dependent on the initial configuration. Cremers *et al.* [Cremers *et al.*, 2006b] treat the case in which there is a number of distinct classes of object in the image by segmenting the image into connected components each of which corresponds to one class. The representation is by distance-function level sets, however, so that overlapping objects are not allowed, while the way in which prior information is included means that only one object can be found in each connected component. There are many tracking methods that use stochastic algorithms, *e.g.* [Isard and Blake, 1998, Kervann and Heitz, 1998, Rathi *et al.*, 2007], but those that deal with multiple objects mostly do so by using the distance-function level set representation. Storvik [Storvik, 1994] uses a Markov Chain Monte Carlo (MCMC) algorithm to minimize an active contour energy, but considers only simply-connected objects; the algorithm makes only local changes to the contour at each iteration. Juan *et al.* [Juan *et al.*, 2006] use stochastic partial differential equation (SPDE) techniques for optimization, but again the stochastic element is limited to small changes to the contour. In contrast, in this paper, although the possible forms of a single-object are limited by adaptation to the data, multiple objects can be created and destroyed at each iteration. Tu *et al.* [Tu *et al.*, 2002] use data-driven reversible jump Markov Chain Monte Carlo (RJMCMC) dynamics to solve a problem of general purpose image segmentation, but our work differs, first, in addressing a specific problem rather than general possibilities; and second, in using multiple birth-and-death rather than RJMCMC dynamics. The advantage of multiple birth-and-death dynamics is their faster convergence, due to the fact that at each iteration, several objects are simultaneously added to the current configuration without any rejection probability.

The remainder of the chapter is organized as follows. In section 6.1, we describe the single-object space and the single-object terms in the energy. In section 6.2, we describe the multiple-object space and the full energy, as well as the sampling algorithm. In section 6.3, we describe experimental results validating the algorithm.

6.1 Single-object space

As mentioned, the single-object space will not be determined *a priori* to consist of arbitrary 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 $\gamma : [0, 2\pi] \to V \subset \mathbb{R}^2$ lying in the image domain V, and we suppose that we are given an energy functional E defined on a space Γ of these curves (with appropriate restrictions to ensure everything is well-defined). This energy functional will depend on the image data also. In this model, it will take the form of a classical active contour energy, which will be detailed below.

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 $\tilde{\cdot} : \Gamma \to \Gamma$ takes every curve to the local minimum in whose basin of attraction it lies. Now define the space $\mathcal{C}$ to be a set of circles lying in the image domain, with radii $r_0 \in [r_{\min}, r_{\max}]$, parametrized by arc-length. The single-object space we consider is $\Gamma_o = \tilde{\mathcal{C}}$. The objects are thus locally adapted to the data and $\tilde{\mathcal{C}}$ consists of a subset of local minima of E obtained by gradient descent from circles $\gamma \in \mathcal{C}$. The dimension of the single-object space is still small however: if we fix the centre $x_0 \in V$ of the circle in $\mathcal{C}$, *i.e.* the 'point' in the marked point process, the 'mark' is one-dimensional, being equivalent to the circle's radius .

6.1.1 Single-object energy

To define the space Γ_o, we define the energy E as a sum of two terms:

$$E(\gamma) = E_{curve}(\gamma) + E_{image}(\gamma) \ , \tag{6.1}$$

where $E_{image}(\gamma)$, is an image term (also called the data term), which relates γ to the image, and $E_{curve}(\gamma)$, is a prior term, which favours boundary smoothness and a uniform parametrization of the curve.

The image energy term is defined as a weighted sum as follows:

$$E_{image}(\gamma) \quad = \quad \lambda_g E_{grad}(\gamma) + \lambda_G E_{gauss}(\gamma) \ ,$$

with

$$E_{\text{grad}}(\gamma) = \int_{[0,2\pi]} dt \; n(t) \cdot \nabla I(\gamma(t)) \tag{6.2}$$

and

$$E_{\text{gauss}}(\gamma) = \int_{R(\gamma)} d^2\mathrm{x} \; (G(\mathrm{x}) - \bar{G}(\mathrm{x})) \ , \tag{6.3}$$

where $n(t)$ is the (unnormalized) outward normal to the curve; I is the image; $G(\mathrm{x}) = \frac{(I(\mathrm{x})-\mu)^2}{2\sigma^2}$ and $\bar{G}(\mathrm{x}) = \frac{(I(\mathrm{x})-\bar{\mu})^2}{2\bar{\sigma}^2}$; 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 when γ is well-adapted to the data. The parameters μ, σ, and $\bar{\mu}$, $\bar{\sigma}$, are learned from examples of object and background.

E_{curve} is defined in the following way:

$$E_{curve}(\gamma) = \int_{[0,2\pi]} dt\, |\dot{\gamma}(t)|^2 \quad , \tag{6.4}$$

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

In order to perform gradient descent method to obtain a curve $\tilde{\gamma} \in \Gamma_0$, we have to compute the functional derivative of E:

$$\frac{\delta E}{\delta \gamma(t)} = \left(\frac{\delta E}{\delta \gamma^x(t)}, \frac{\delta E}{\delta \gamma^y(t)} \right) \tag{6.5}$$

For our computation, we define the curves $\gamma(t)$ to be traversed in the clockwise direction, so that the outward normal to the curve can be written as follows:

$$\begin{aligned} n^x(\gamma(t)) &= -\dot{\gamma}^y(t) \\ n^y(\gamma(t)) &= \dot{\gamma}^x(t) \end{aligned}$$

After that, we consider a vector field $v(\gamma)$ defined as:

$$\begin{aligned} v^x(\gamma(t)) &= -\partial_y I(\gamma(t)), \\ v^y(\gamma(t)) &= \partial_x I(\gamma(t)), \end{aligned}$$

i.e. $v(\gamma) = \epsilon \nabla I(\gamma)$, where $\epsilon = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix} \right)$ is a rotation matrix. Thus, we can write:

$$E_{grad}(\gamma) = \int_{[0,2\pi]} dt\, n(\gamma(t)) \cdot \nabla I(\gamma(t)) = \int_{[0,2\pi]} dt\, \dot{\gamma}(t) \cdot v(\gamma(t)) \quad . \tag{6.6}$$

The components of the E_{grad} derivative then take the form:

$$\frac{\delta E_{grad}}{\delta \gamma^x(t)} = -\dot{\gamma}^y(t)[\partial_y v^x(\gamma(t)) - \partial_x v^y(\gamma(t))] \quad ,$$

$$\frac{\delta E_{grad}}{\delta \gamma^y(t)} = -\dot{\gamma}^x(t)[\partial_x v^y(\gamma(t)) - \partial_y v^x(\gamma(t))] \quad ,$$

which give

$$\frac{\delta E_{grad}}{\delta \gamma^i(t)} = n^i(t)(\nabla \times v)(\gamma(t)) \quad . \tag{6.7}$$

For the calculation details, please refer to the appendix, section B.1. Taking into account that $v(\gamma) = \epsilon \nabla I(\gamma)$, the expression (6.7) can be finally written as follows:

$$\begin{aligned} \frac{\delta E_{grad}}{\delta \gamma^i(t)} &= n^i(t)\,(\nabla \times \epsilon\, \nabla I)\,(\gamma(t)) \\ &= n^i(t)\,(\nabla \cdot \nabla I)\,(\gamma(t)) \\ &= n^i(t)\,\nabla^2 I(\gamma(t)) \end{aligned}$$

To calculate, now, the derivative of the Gaussian energy term E_{gauss}, it is possible to represent it in the same way as E_{grad}. In order to compute the derivative, we consider the functions $U(\mathrm{x})$ and $h(\mathrm{x})$ defined as $h(\mathrm{x}) = \nabla \cdot U(\mathrm{x}) = (G - \bar{G})(\mathrm{x})$, so that (6.3), using Green's theorem can be written:

$$\begin{aligned} E_{gauss}(\gamma(t)) &= \int_{R(\gamma)} d^2\mathrm{x}\,(G(\mathrm{x}) - \bar{G}(\mathrm{x})) \\ &= \int_{R(\gamma)} d^2\mathrm{x}\, h(\mathrm{x}) \\ &= \int_{R(\gamma)} d^2\mathrm{x}\, \nabla \cdot U(\mathrm{x}) \\ &= \int_{\partial R(\gamma)} dt\, n(\gamma(t)) \cdot U(\gamma(t)) \end{aligned} \tag{6.8}$$

Expression (6.8), in its turn, can be written using a curve derivative:

$$\int_{\gamma(t)} dt\, n(\gamma(t)) \cdot U(\gamma(t)) = \int_{\gamma(t)} dt\, \dot{\gamma}(t) \cdot v(\gamma(t)) \quad ,$$

where v is now defined as a rotated version of U:

$$\begin{aligned} v^x(\gamma(t)) &= -\,U^y(\gamma(t)) \\ v^y(\gamma(t)) &= U^x(\gamma(t)) \quad . \end{aligned}$$

The derivative of E_{gauss}, thus, takes the form:

$$\begin{aligned} \frac{\delta E_{gauss}}{\delta \gamma^i(t)} &= n^i(t)(\nabla \times v)(\gamma(t)) \\ &= n^i(t)\,(\nabla \times \epsilon\, U)\,(\gamma(t)) \\ &= n^i(t)\,(\nabla \cdot U)\,(\gamma(t)) \\ &= n^i(t)\,(G - \bar{G})(\gamma(t)) \quad . \end{aligned}$$

Finally, the computation of the derivative of E_{curve} (6.4) is as follows. We vary, first, E_{curve} by $\delta_{\gamma(t)}$:

$$\begin{aligned}
\delta_{\gamma(t)} E_{curve} &= E_{curve}(\gamma + \delta\gamma) - E(\gamma) \\
&= \int_{[0,2\pi]} dt \, \left(\dot{\gamma}^2(t) + 2\dot{\gamma}(t)\delta\dot{\gamma}(t) + O(\delta^2\dot{\gamma}(t)) \right) - \int_{[0,2\pi]} dt \, |\dot{\gamma}(t)|^2 \qquad (6.9) \\
&= 2 \int_{[0,2\pi]} dt \, \dot{\gamma}(t)\delta\dot{\gamma}(t) \qquad (6.10) \\
&= 2 \left(\dot{\gamma}(t)\delta\gamma(t)|_0^{2\pi} - \int_{[0,2\pi]} dt \, \ddot{\gamma}(t)\delta\gamma(t) \right) \\
&= -2 \int_{[0,2\pi]} dt \, \ddot{\gamma}(t)\delta\gamma(t) \ , \qquad (6.11)
\end{aligned}$$

Then the derivative takes the form:

$$\frac{\delta E_{curve}}{\delta \gamma^i(t)} = -2\ddot{\gamma}^i(t) \ . \qquad (6.12)$$

Figures 6.1 and 6.2 show gradient descent experiments using the energy E, where the procedure of gradient descent is as follows. The curves considered are represented by a sequence of points in $\mathbb{R}^2$ defined to correspond to descrete parameter values $t_n = 2\pi n/N$ for $n \in \{0, \ldots, (N-1)\}$. We define the set of circles $\mathcal{C}$ lying in the image, with radii in the range $[r_{\min}, r_{\max}]$ and with centres at the image pixels. The circles in $\mathcal{C}$ are assumed to be arc-length parametrized, therefore the points are equally spaced. A curve is initialized by a circle lying in the image with some radius from $[r_{\min}, r_{\max}]$. It is then evolved under the gradient field (6.5), where the derivative (6.12) of the curve smoothness energy term also controls the curve parametrization: in the implementation, the discrete version of the derivative is given by

$$\frac{\delta E_{curve}}{\delta \gamma(t_n)} = -2 \frac{\gamma(t_{n+1}) - 2\gamma(t_n) + \gamma(t_{n-1})}{(\Delta t)^2},$$

where $\Delta t = t_n - t_{n-1} = 2\pi/N$, which encourages the points to be equispaced along the curve. In detail, the gradient descent procedure is as described in algorithm 6.1.

Algorithm 6.1 Gradient descent under the gradient field of energy E

1. Initialize an initial curve as a circle $\gamma \in \mathcal{C}$ lying in the image domain; the parameters of E; discrete step δ^E for gradient descent; ϵ for stop condition; and $k = 0$ the iteration number;
2. Compute $\frac{\delta E_{curve}}{\delta \gamma^{(k)}(t_n)} = -2\frac{\gamma^{(k)}(t_{n+1})-2\gamma^{(k)}(t_n)+\gamma^{(k)}(t_{n-1})}{(\Delta t)^2} + \lambda_g n^{(k)}(t_n) \cdot \nabla^2 I(\gamma^{(k)}(t_n)) + \lambda_G n^{(k)}(t_n) \cdot (G - \bar{G})(\gamma^{(k)}(t_n)), n = 0, \ldots, N-1$;
3. Update the current curve: $\gamma^{(k+1)}(t_n) = \gamma^{(k)}(t_n) - \delta^E \frac{\delta E_{curve}}{\delta \gamma^{(k)}(t_n)}, n = 0, \ldots, N-1$;
4. Compute $\Delta\gamma^{(k+1)} = \left\|\gamma^{(k+1)} - \gamma^{(k)}\right\|$;
5. Stop if $\left\|\Delta\gamma^{(k+1)} - \Delta\gamma^{(k)}\right\| < \epsilon$, where $\|\cdot\|$ is a Euclidean norm ($\Delta\gamma^{(0)}$ is equal to some large number, such that $\left\|\Delta\gamma^{(1)} - \Delta\gamma^{(0)}\right\| > \epsilon$); if not, then go to step 2.

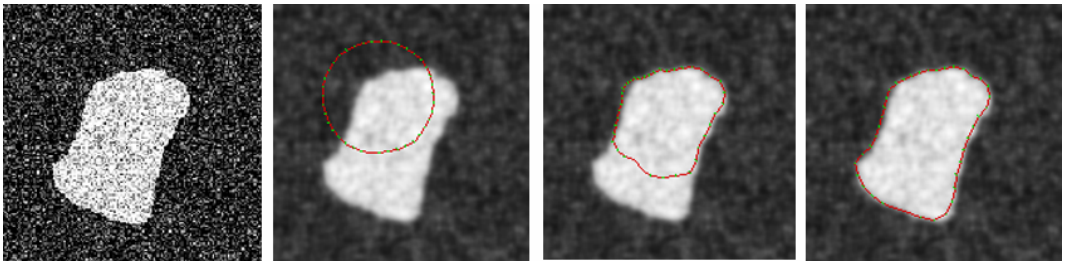

Figure 6.1: Synthetic image example: gradient descent driven by E. On the left is the original image, binary with added noise. The other three images show the contour evolution. The background has been blurred for display purposes.

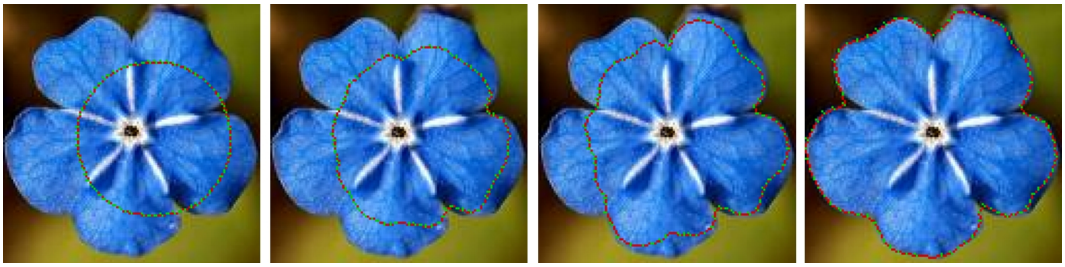

Figure 6.2: Real image example: gradient descent driven by E. Only one band of the colour image was used.

6.2 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:

$$\Omega_{\Gamma_o} = \bigcup_{n=0}^{\infty} \left[\Gamma_o^n / S_n \right] \ , \tag{6.13}$$

where S_n indicates the symmetric group of n elements acting on the components of the product. Set union and difference are defined for appropriate pairs of elements of Ω_{Γ_o}. The map $\tilde{\cdot}$ extends to a map from $\Omega_{\mathcal{C}}$ (the exponential of $\mathcal{C}$) to Ω_{Γ_o}. Elements of $\Omega_{\mathcal{C}}$ will be denoted ω, which is an non-ordered set of components $\omega_i \in \mathcal{C}$: $\omega = \{\omega_i\}$. Every ω_i can be represented by its center $x_0^{\omega_i}$ and its radius $r_0^{\omega_i}$ so that, $\omega_i = (x_0^{\omega_i}, r_0^{\omega_i})$.

6.2.1 Energy

Given a (real, bounded below) function $H(\omega)$ on $\Omega_{\mathcal{C}}$, we define the Gibbs distribution μ_β in terms of the density $p(\omega) = \frac{d\mu_\beta}{d\mu}(\omega)$ w.r.t. the reference measure μ on the space $\Omega_{\mathcal{C}}$ of configurations of circles $\omega = (x_0^\omega, r_0^\omega)$, being respectively the centres x_0^ω of circles in the configuration ω and their radii $r_0^\omega \in [r_{\min}, r_{\max}]^{|\omega|}$, where $|\omega|$ is the number of elements in ω, so that, defined as

$$d\mu(\omega) = d\lambda(x_0^\omega)\, d\nu(r_0^\omega) \ , \tag{6.14}$$

where $d\lambda(x_0^\omega)$ is a Lebesgue-Poisson measure on the space of configurations of centres, and $d\nu(r_0^\omega) = \prod_{x_0 \in x_0^\omega} d\nu(r_0) = \prod_{x_0 \in x_0^\omega} \frac{dr_0}{l}$, $l = |r_{max} - r_{min}|$, is a conditional on the number of objects (depending on the centres in the given configuration ω elements) measure defined as a product of measures ν on the space of marks. The Gibbs density $p(\omega)$ with respect to μ is then defined as

$$p(\omega) = \frac{z^{|\omega|}}{Z_\beta} \exp\{-\beta H(\omega)\} \ , \tag{6.15}$$

with parameters $\beta > 0$, $z > 0$, and where the normalization constant Z_β is given by :

$$\begin{aligned} Z_\beta &= \int_{\Omega_{\mathcal{C}}} d\mu(\omega)\, z^{|\omega|} \ \exp\{-\beta H(\omega)\} \\ &= 1 + \sum_{n=1}^{\infty} \frac{z^n}{n!} \int_{\mathcal{C}^n} d\mu(\omega) \ \exp\{-\beta H(\omega)\} \ . \end{aligned} \tag{6.16}$$

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

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

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

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

The term H_2 is the interaction term, which controls the relation between objects, in particular discouraging overlaps. It is defined as

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

where

$$\delta_\epsilon(\omega_i, \omega_j) = \begin{cases} \infty, & |x_0^i - x_0^j| \leq \epsilon; \\ 0, & \text{otherwise,} \end{cases} \tag{6.17}$$

A is the area function 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-objects.

Computing the intersection area of two objects with complex shape is quite a non-trivial problem when compared to that of simply-shaped objects, *e.g.* circles, ellipses or rectangles. We approximate it by the number of pixels that belong to the interiors of both objects. To compute it, we need to obtain the interior pixels of the objects. We developed an algorithm for computing the interior area of a discretized object contour represented by the chain of points corresponding to its vertices with known coordinates in the image. The idea is as follows. Firstly, we reconstruct the contour by connecting its vertices with segments. The segments are obtained using Bresenham's algorithm [Bresenham, 1965]. Then, we build up the inside contour. Once the two contours are obtained, the interior of the object can be obtained as a set of pixels which are inside the inside contour and are not separated by the contour itself plus the pixels belonging to both contours. Figure 6.3 illustrates the contour with its inside contour, where the vertices of the contour are the pixels surrounded by little circles for display purposes.

The procedure is detailed in algorithm 6.2.

To avoid an infinite loop at step 1, the contour is defined as a badly-defined curve if I is not found in 50 tries, and is rejected as an impossible object. This can happen when the points approximating the curve are situated too close to each other, or when the curve is very irregular and sharp. However, in practice, this happens extremely seldom, since in our experiments the objects are modeled so that the uniform parametrization is favoured and the objects detected are big enough to prevent curve collapsing.

6.2.2 Model simulation

In order to simulate the proposed model for finding the optimal configuration of objects in the image, we consider a multiple birth-and-death dynamics, defined in $V \subset \mathbb{R}^2$ by its generator [Descombes *et al.*, 2009] in the space of bounded measurable functions in $\Omega_{\mathcal{C}}$:

Algorithm 6.2 Reconstruction of the interior of a contour

1. Find an inside point that does not belong to the contour and is situated just next to it; we note it I (it is needed to build up the inside contour). This pixel could be found by scanning a rectangle containing the contour. We use the following random method:
   (a) Reconstruct the contour connecting the vertices with segments;
   (b) Randomly determine a straight line which does not intersect with the contour;
   (c) Determine the vertex closest to the line; we denote this vertex $v1$ in the figure;
   (d) Determine a barycentre denoted B of vertices $v1$, $v2$, $v3$, where the vertices $v2$ and $v3$ are the previous and the following neighbours of $v1$;
   (e) If the pixel next to $v1$ on the line connecting it with B, does not belong to the contour, then it belongs to the inside contour and it is the pixel I that we look for; if not, then go to the step 1.(a).
2. Reconstruct the inside contour beginning from the obtained pixel I, in such a way, that every pixel composing it is connected to the previous one through the 4-connectivity neighbourhood relation as shown in figure 6.3. In order to do it, we used a fast algorithm presented in [Khudeev, 2005]. The idea is to begin the reconstruction of the inside contour at some inside pixel I just next to the contour. Then, to move along the contour until being back to the first pixel I of the inside contour.
3. Find the interior of the contour:
   (a) Order the pixels of both the contour itself and the inside contour w.r.t. the x coordinate and then w.r.t the y coordinate in the image;
   (b) For each row of the rectangle containing the contour, defined by the maximum and minimum coordinates x and y of the contour, mark every pixel between two inside contour pixels that are not separated by a contour pixel as an inside pixel.

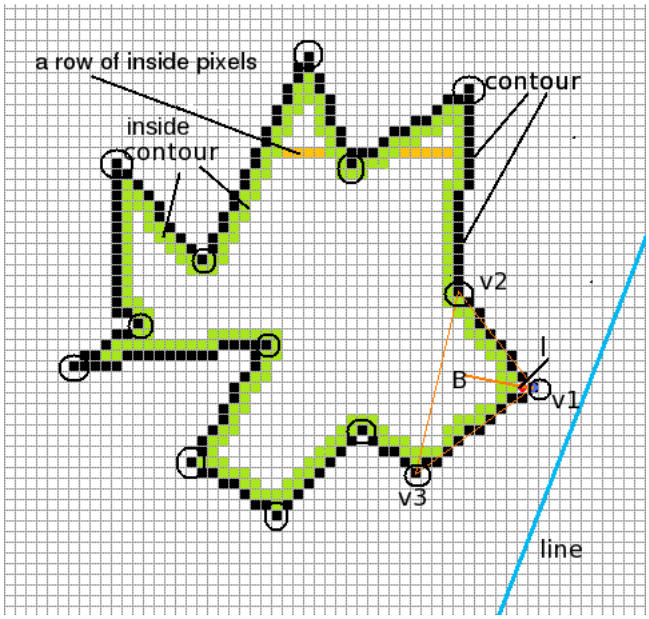

Figure 6.3: An illustration to algorithm 6.2 for computing the internal area of a contour.

$$(L_\beta f)(\omega) = \sum_{\omega_i \in \omega} e^{\beta \Delta_i H(\omega)} (f(\omega \setminus \omega_i) - f(\omega)) + z \int_{\mathcal{C} \setminus \gamma} f(\omega \cup \gamma) - f(\omega)\, dx_0\, d\nu(r_0), \quad (6.18)$$

where the energy difference for a configuration ω with and without an element ω_i is given by

$$\Delta_i H(\omega) = H(\omega) - H(\omega \setminus \omega_i). \quad (6.19)$$

The intensity $b(\omega, \gamma)$ of adding element $\gamma = (x_0, r_0) \in \mathcal{C}$ to a configuration ω is called birth intensity and is given by:

$$b(\omega, \gamma) = z, \quad (6.20)$$

and the intensity $d(\omega, \omega_i)$ of removing an element ω_i from the configuration, which is called death intensity is defined as:

$$d(\omega, \omega_i) = e^{\beta \Delta_i H(\omega)}. \quad (6.21)$$

Under such a choice of intensities, the detailed balance condition, given here below, holds:

$$\frac{b(\omega \setminus \omega_i)}{d(\omega, \omega_i)} = \frac{p(\omega)}{p(\omega \setminus \omega_i)} = z\, e^{-\beta \Delta_i H(\omega)}.$$

This implies that the corresponding birth-and-death process associated with the stochastic semi-group $T_\beta(t) = e^{tL_\beta}$ is time reversible. Thus, its equilibrium distribution is the Gibbs

stationary measure μ_β defined by (6.15).

As we mentioned in the introduction, the aim of the work described in this section is to lift the restriction of MPP models to the extraction of objects with only simple shapes, keeping, nevertheless, the advantage of the small dimensionality of the space of individual objects, since the inverse would greatly increase the computational complexity of sampling and estimation. Therefore, we will define the process in $\Omega_{\mathcal{C}}$. In the space of arbitrarily-shaped elements locally adapted to the data Ω_{Γ_o}, where $\Gamma_o = \tilde{\mathcal{C}}$, the process is equivalent, the only difference being that the birth intensity is non-uniform. This is due to the fact that under gradient descent two different circles can give the same element $\tilde{\gamma}$, which means that this element has higher probability to be born. However, it does not change the final result, since this affects the reference measure and not the density. As our process is embedded in the annealing scheme, the results will be influenced only in the first steps, when the temperature is high and the influence of the reference measure is high; as the temperature tends to zero, the influence of the density function begins to dominate. For the development of our approach, we refer to the theorems listed in section 6.2.4 below.

6.2.3 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. To perform the sampling process we consider a Markov chain $T_{\beta,\delta}(m), m = 0, 1, 2 \ldots$ in $\Omega_{\mathcal{C}}$ 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. Here, the new objects are given by ω'', and the removed objects are given by $\omega \setminus \omega'$.

Descombes *et al.* demonstrate in [Descombes *et al.*, 2009], that this Markov chain can be considered as an approximation of a continuous-time reversible process $T_\beta(t)$ and converging to it, which, within a logarithmic annealing scheme, guarantees weak convergence to the measure concentrated on the global minima of the energy function $H(\omega)$, where $\omega_i \in \omega$ is a circle with a fixed radius.

For realisation of this continuous-time process, we considered a discrete-time approximation $T_{\beta,\delta}(m)$ of the continuous-time birth-and-death process defined by its generator (6.18). This approximation scheme makes uses of a Markov chain in $\Omega_{\mathcal{C}}$ with multiple birth-and-birth transitions, defined here below. The transitions of 'birth' of a new element γ centred at x_0 from $\Delta x_0 \in V$ and with a radius r_0 from $\Delta r_0 \in [r_{\min}, r_{\max}]$ have the distribution

$$q_\gamma = \begin{cases} \delta z \Delta x_0 \frac{\Delta r_0}{l}, & \text{if } \omega \to \omega \cup \gamma, \\ 1 - \delta z \Delta x_0 \frac{\Delta r_0}{l}, & \text{if } \omega \to \omega. \end{cases} \tag{6.22}$$

The transition probability of 'death' of an element ω_i is given by

$$p_d = \begin{cases} \frac{\delta d_\beta}{1+\delta d_\beta}, & \text{if } \omega \to \omega \setminus \omega_i, \\ \frac{1}{1+\delta d_\beta}, & \text{if } \omega \to \omega, \end{cases} \tag{6.23}$$

where $d_\beta = \delta e^{\beta \Delta_i H(\omega \setminus \omega_i)}$.

The transition operator $P_{\beta,\delta}$ of this Markov chain $T_{\beta,\delta}(m) = P^m_{\beta,\delta}$ is defined as:

$$\begin{aligned} (P_{\beta,\delta} f)(\omega) \quad = \quad & \sum_{\omega' \subseteq \omega} \prod_{\omega'_i \in \omega'} \frac{1}{1+\delta d_\beta} \prod_{\omega_i \in \omega \setminus \omega'} \frac{\delta d_\beta}{1+\delta d_\beta} \\ & \times \frac{1}{\Xi_\delta(\omega')} \sum_{k=0}^{\infty} \int_{\Omega^k_{\mathcal{C}}(\omega')} \frac{(z\delta)^k}{k!} \\ & \times f(\omega' \cup \gamma'_1 \cup ... \cup \gamma'_k) dx^1_0 ... dx^k_0 d\nu^1(r_0) \dots d\nu^k(r_0), \end{aligned} \tag{6.24}$$

where $\Xi_\delta(\omega') = \Xi_\delta(x^{\omega'}_0, r^{\omega'}_0, z, \delta)$ is a normalizing factor for the conditional measure under the given configuration ω'.

6.2.4 Reference theorems

In this section we formulate the theoretical results from [Descombes *et al.*, 2009] that form the basis of our algorithm. These results were obtained in the case of circles with a fixed radius r_0. In our case, when a set of radius is finite the generalisation of these results is straightforward.

Let $\Omega^{r_0}_{\mathcal{C}}$ be the set of $\omega \in \Omega_{\mathcal{C}}$ whose elements $\gamma \in \mathcal{C}$ have a fixed radius r_0. Denote $\bar{H} = \min_{\omega \in \Omega^{r_0}_{\mathcal{C}}} H(\omega)$ and

$$H^{min} = \left\{ \omega \in \Omega^{r_0}_{\mathcal{C}} : H(\omega) = \bar{H} \right\} \tag{6.25}$$

the set of all elements in $\Omega^{r_0}_{\mathcal{C}}$ giving the global minimum $\bar{H}$ of $H(\omega)$. Then, H^{min} can be represented as a union

$$H^{min} = \bigcup_{n=0}^{\infty} H^{min}_n \quad , \tag{6.26}$$

where H^{min}_n is a set of configurations from H^{min} containing n elements. The following theorem states that the Gibbs distribution μ_β converges to the measure concentrated on the global minima of H with the minimum number of objects n_0.

Theorem 6.4. *Let $n_0 \in [0, ...N]$ be the minimal index for which the set H^{min}_n is not empty. Then Gibbs distribution μ_β converges weakly as $\beta \to \infty$ to the distribution μ_∞ on $\Omega^{r_0}_{\mathcal{C}}$ of the*

form

$$\begin{aligned} \mu_\infty &= \textstyle\sum_{\omega\in H_{n_0}^{min}} C_\omega \delta_\omega, \ \ if\ n_0 > 0, \ and \\ \mu_\infty &= \delta\{\emptyset\}, \ \ if\ n_0 = 0. \end{aligned} \tag{6.27}$$

Here δ_ω is a unit measure concentrated on the configuration ω, and the coefficients C_ω satisfy the equality

$$\sum_{\omega\in H_{n_0}^{min}} C_\omega = 1 \ \ . \tag{6.28}$$

The following theorem guarantees uniform convergence to the measure concentrated on the global minima of the energy H.

Let $\mathcal{B}(\mu)$ be a family of measures η on $\Omega_{\mathcal{C}}^{r_0}$ with a bounded density w.r.t. the reference measure μ, and hence w.r.t. the Gibbs measure μ_β, therefore $p_\eta(\omega) = \frac{d\eta}{d\mu_\beta}$ on $\Omega_{\mathcal{C}}^{r_0}$ is a bounded density of the measure η w.r.t. the Gibbs measure μ_β.

Theorem 6.5. *Let F be a bounded function in $\Omega_{\mathcal{C}}^{r_0}$ and an initial measure $\eta \in \mathcal{B}(\mu)$. Then, under the condition that*

$$\delta e^{\beta b} < const, \tag{6.29}$$

where $b = \sup_{\omega\in\Omega_{\mathcal{C}}^{r_0}} \sup_{\gamma\in\omega} \Delta_i H(\omega)$ we have

$$\lim_{\beta\to\infty,\ t\to\infty,\ \delta\to 0} \langle F\rangle_{S_{\beta,\delta([\frac{t}{\delta}])}\eta} = \langle F\rangle_{\mu_\infty}\,. \tag{6.30}$$

Here

$$\langle F\rangle_{S_{\beta,\delta([\frac{t}{\delta}])}\eta} = \left\langle S_{\beta,\delta([\frac{t}{\delta}])}\eta, F\right\rangle = \left\langle \eta, T_{\beta,\delta([\frac{t}{\delta}])}F\right\rangle = (p_\eta, T_{\beta,\delta([\frac{t}{\delta}])}F)_{\mu_\beta}, \tag{6.31}$$

where $S_{\beta,\delta([\frac{t}{\delta}])}$ indicating the semi-group adjoint to $T_{\beta,\delta([\frac{t}{\delta}])}$ acting on measures.

For the details and the proof of the theorems, the reader can refer to [Descombes *et al.*, 2009].

6.2.5 Algorithm description

We define $\mathcal{C}$ as the set of circles lying in the image domain V, with radii in the range $[r_{\min}, r_{\max}] \cap \mathbb{N}$ and with centres at the image pixels. The curves are represented by a chain of points in $\mathbb{R}^2$ defined to correspond to discrete parameter values $t_n = 2\pi n/N$ for $n \in \{0, \ldots, (N-1)\}$. The circles in $\mathcal{C}$ are assumed to have arc-length parametrization, 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 $\Delta_i H(\omega) = H(\omega \backslash \omega_i) - H(\omega)$. In more detail, the algorithm is as follows:

1. **Initialization**

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

2. **Birth**

   (a) Sample a configuration of circles with radii uniformly distributed on $[r_{\min}, r_{\max}] \cap \mathbb{N}$, 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 $\omega \in \Omega_{\mathcal{C}}$;

   (b) Evolve every circle in ω using gradient descent, with gradient field given by equation (6.5), until convergence, producing configuration $\tilde{\omega} \in \Omega_{\Gamma_o}$; and add the obtained elements to the current configuration;

3. **Death**

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

   (b) Remove each component ω_i from the current configuration with probability

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

   where

   $$d_\beta(\omega_i, \omega) = e^{-\beta \Delta_i H(\omega)} \quad ; \tag{6.32}$$

4. **Termination**

   If all the components added in the birth step and only these, are removed in the following death step, then stop; if not, then increase the inverse temperature β by a factor $\Delta\beta$ and decrease the time step δ by a factor $\Delta\delta$, and go to the birth step.

6.3 Experimental results

In this section we present several results obtained using the above model. Experimental results are presented on a synthetic binary noisy image and then on a real image of flowers

of quite a complex shape but without noise in the background, and finally, on a real noisy CIR image of 3 cm resolution representing the tops of tree crowns with complex shapes. We present the original image and the obtained extracted objects. On several images, value of the terms in the energy are displayed. We give the computation time to give an idea of resources needed. The algorithm was tested on a 2.16 GHz processor.

There is a set of parameters to be fixed before testing the algorithm. For our experiments we use a geometric annealing schedule for the inverse temperature β and a time step δ. The time-discretisation step update $\Delta\delta = 0.99$ and the inverse temperature update $\Delta\beta = \frac{1}{0.993}$ were fixed to the given values and were the same for all the experiments. The other parameters depend on the image and the objects to be found. The minimum and maximum radius values r_{min}, r_{max} that determine the objects marks, *i.e.* the radii of initial circles, are calibrated knowing the image resolution and the approximate size of the objects. The parameters of the data energy term, $\sigma, \mu, \bar{\sigma}, \bar{\mu}$ are learned using a fragment of image with object and background examples. The weighting parameters are learned from experiments.

Figure 6.4 shows the result of an experiment on a synthetic binary image with additive white Gaussian noise. Due to the simple shape of the objects, the computational time is small (3 mn), since the gradient descent algorithm converges fast to a local minimum of the data energy giving thus objects locally adapted to the data. This image contains distinct objects and two overlapping objects in the left corner. The objects have quite circular shapes except four of them: the two overlapping objects, and the two next to them in the top of the image. All the extracted objects are detected and well delineated, except the two in the top middle part of the image. The loss in delineation precision is due to the fact that the smoothness term that also controls the length of the curve, was multiplied by a large weight in order to favour the detection of circular objects. This was done in order to detect the two overlapping objects with two distinct elements. Otherwise, they could be considered as one more complex object that would be accurately delineated, but two distinct objects would be enveloped by one curve of elongated shape, cf figure 6.5.

The result of another experiment on one band of a real colour image is shown in figure 6.6. We aimed at extracting the flowers without detecting their stems. The most discriminative band between the flower and its stem is the red, which we chose for learning the parameters for the Gaussian image model. Since the radiometry of the stems and flowers is close compared to the difference with the background, two flowers were detected together connected by a stem, as shown in figure 6.7. To avoid this effect, the smoothness term was increased again, since the flower heads remain relatively regular, but as we can see some small details of the flower petals were smoothed and lost as expected.

The last experiments that we present were done on a fragment (345x340 pixels) of a real colour infra-red aerial image of a scene composed of tree crowns. The image viewpoint is close to the nadir, *i.e.* the tree crowns are seen from above. We chose this part of the image

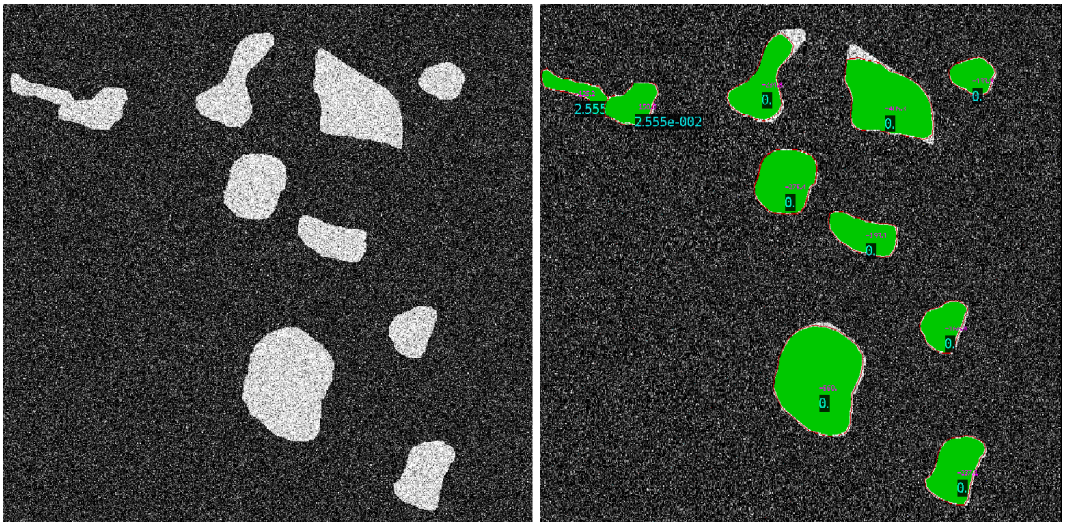

Figure 6.4: Left: original synthetic binary image with added noise. Right: final object configuration. The numbers in the interiors of the curves show the value of H_2 (black background), demonstrating that there is no degeneracy in the solution, and the value of H_1.

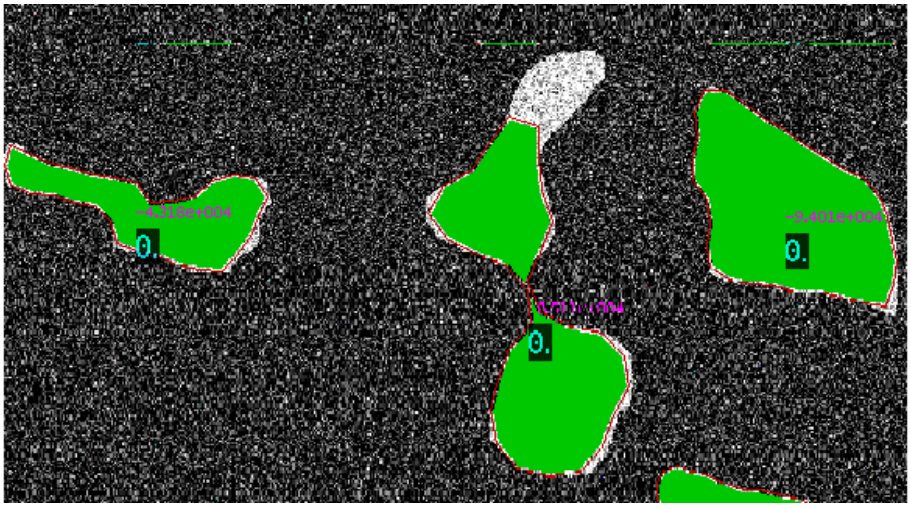

Figure 6.5: Illustration of the effect of delineation of two objects by one curve, when the prior curve term E_{curve} is dominated by the image term E_{image}.

for our experiments in order to test the approach on a scene of objects that have quite regular as well as very spiky shapes, and with some overlapping objects. In figures 6.8 and 6.9 we

Figure 6.6: Left: original real image. Right: final object configuration. The numbers in the interiors of the curves show the value of H_2 (black background), demonstrating that there is no degeneracy in the solution, and the value of H_1.

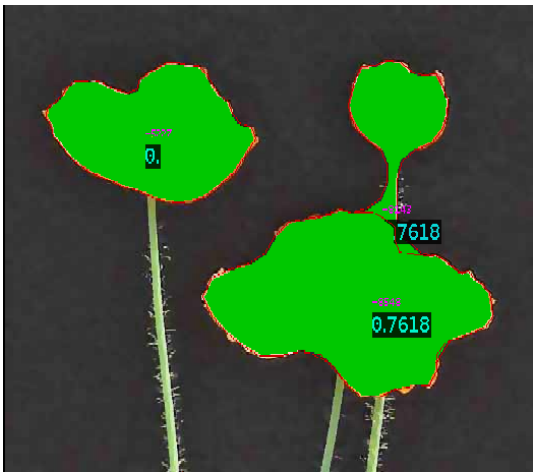

Figure 6.7: Illustration of two detected flowers connected by their stems, when E_{curve} is not weighted enough, allowing the curve shape to be far from a circle. The numbers in the interiors of the curves show the value of H_2 (black background), demonstrating that there is no degeneracy in the solution, and the value of H_1.

show the configurations resulting from two tests: the first test is done without favouring the smoothness term; the second test is done with the smoothness term favoured. The interior of the detected curves is transparent to show the object overlaps. For this image, the method

for separating the objects by increasing the smoothness term deals with the problem only to some degree as we can see from the result. Two overlapping crowns situated in the bottom of the image are detected by two distinct objects that overlap, (as the weight in front of the interaction term varies, the degree of overlap changes), whereas in the left and middle part of the image we can observe the inverse effect. As in the previous experiments, the precision in the delineation of the objects decreases as we favour E_{curve}, which also controls the curve length, thus preventing the curve from being stretched much in different directions. As result, the boundaries and thus the details of the curves are greatly smoothed, *e.g.* the peaks of the spruce trees (the trees with spiky shapes). The computational time for this image was approximately 15 mn for both tests.

A set of experiments of different levels of complexity was done to test the model. From the obtained results, we can conclude that 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 solution that we propose to deal with this problem is to incorporate more specific prior information into the energy E in order to detect objects in scenes of high complexity containing overlapping objects, in particular, in forests, cf figure 6.8, with a resolution extending almost to leaf level. We define a method for incorporating prior information about the shape of objects into the single-object model. This method is presented in the following chapter.

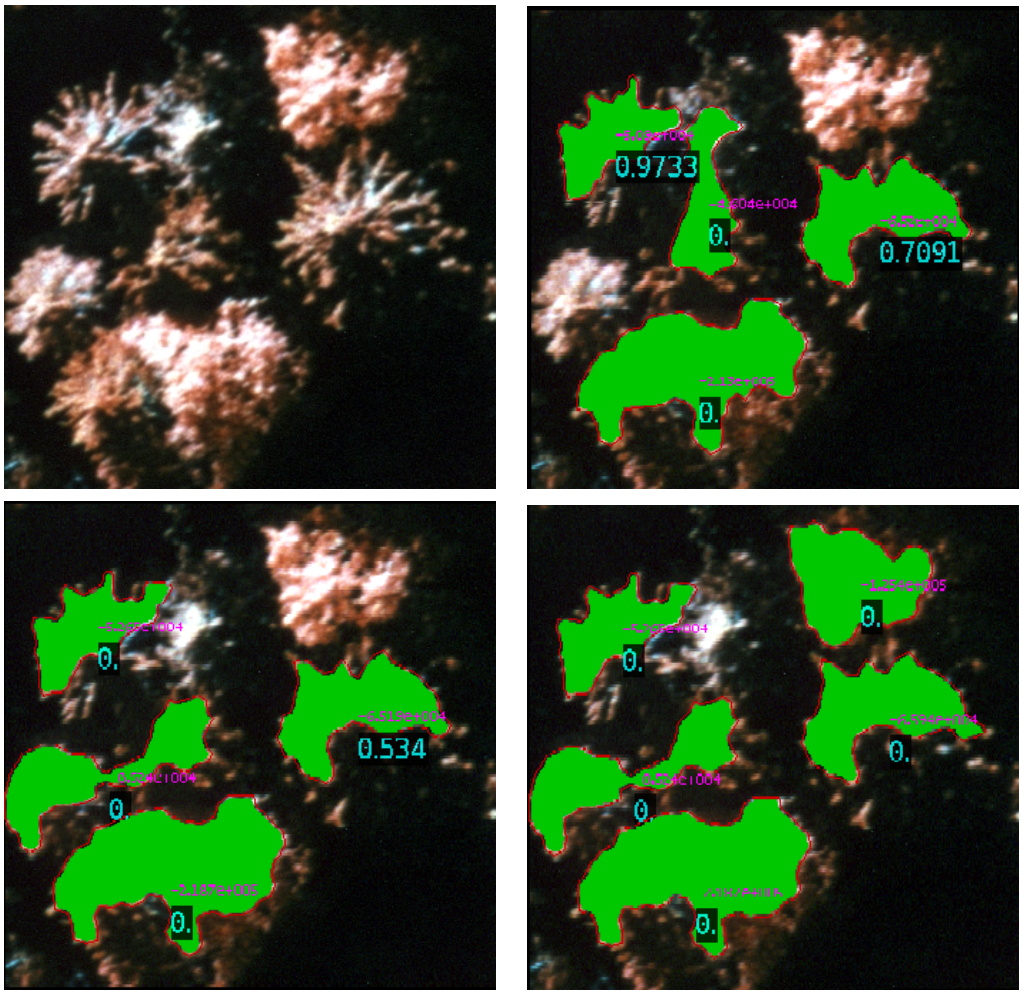

Figure 6.8: Top, left: original CIR image of tree crowns, © CBA. Bottom, right: final configuration obtained. The intermediate images show the configuration evolution during the birth-and-death process. The numbers on the crowns show the interaction term value H_2 (black background) and the data term value H_1.

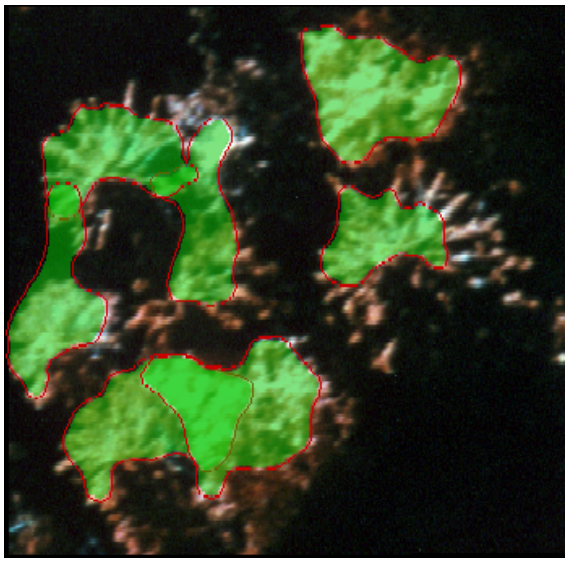

Figure 6.9: Final configuration, when E_{curve} is favoured in order to avoid detecting the overlapping crowns as one object.

Chapter 7

Marked Point Process with shape prior for multiple arbitrarily-shaped object extraction

7.1 Introduction

In the previous chapter, we presented a marked point process model of multiple arbitrarily-shaped objects: the restriction imposed by the simplified nature of the shape of objects involved was lifted without increasing the dimension of the single-object space [Kulikova *et al.*, 2010]. Every 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. The single-object model considered included only weak shape information, thus allowing the delineation of the distinct objects of arbitrary and smooth enough boundaries, whereas overlapping objects with similar radiometric characteristics were detected as a single object. A possible way to increase to some degree the accuracy in the number of segmented objects consisted in favouring the energy smoothness term, but this also decreased the precision of object delineation.

In this chapter, we define a method for incorporating strong prior information about the shape of the objects sought into the single-object model, in order to deal with overlapping objects with complex shapes without losing their geometric details and without significantly increasing the computational complexity of estimation.

The extended MPP model including a strong shape prior can also be viewed as the active contour methodology for an unknown *a priori* number of objects, where prior knowledge about object shape is incorporated as well. Much work has already been devoted to the active contour approach. Some of this work, *e.g.* [Caselles *et al.*, 1997], includes only

weak shape information, essentially smoothness, but can in principle detect multiple (although not overlapping) objects using the level set representation [Osher and Fedkiw, 2003, Sethian, 1999]. Other work includes much stronger prior information about shape. Cremers *et al.* , for example, incorporate statistical prior information about the shape of the object to be extracted in a modified Mumford-Shah functional, the segmenting curve being represented by a closed parametrized spline curve [Cremers *et al.*, 2006c]. A statistical model is then built on a set of training shapes using a Gaussian distribution on the spline control point vectors, such that no projection on the subspace of training sample deformations is needed, in contrast with the approach described in [Leventon and Grimson, 2000]. Leventon *et al.* propose to include a statistical shape prior into the distance-function level set representation of active contours. A statistical shape model is constructed using the distribution over a set of training shapes. An initial curve is embedded as the zero level set of a higher dimensional surface, which is evolved so that its zero level set fits in the limit the boundary of the target object. In [Rousson and Paragios, 2002, Cremers *et al.*, 2006b] some other methods incorporating strong shape prior are defined using a level set framework for image segmentation. [Joshi and Srivastava, 2009] describes a way for incorporating prior shape knowledge into an active contour model, where the prior statistical shape model is built on the tangent space of non-linear shape manifold. The limitations of these methods are the complexity of their application as well as the difficulty of treating an unknown number of distinct or overlapping objects (and in practice only single objects are treated).

Below, we describe a process for the incorporation of a shape prior information into the MPP model defined in the previous chapter. The main idea of the model remains the same, but the single-object model is extended. We give the details of the incorporated modifications in section 7.2. We define, first, the single-object space, where a single object is defined by the data and the shape prior information, with a strong shape prior included using a representation of a curve by its centre and a radial variation around a circle, the shapes being therefore 'star domains'. Then, in section 7.3, we present a set of experiments on real 3cm/pixel resolution aerial images of forest. Finally, we summarize in section 7.4, where we also give a comparative example of the application to tree crown extraction of MPP models with different shape priors included: from simply-shaped objects to arbitrarily-shaped ones with strong shape prior information incorporated.

7.2 Single-object energy

We model individual object boundaries as closed planar curves $\gamma : [0, 2\pi] \to V \subset \mathbb{R}^2$ lying in the image domain V. The set of closed curves we consider here consists of 'star domains' parameterized by $(x_0, \delta r(t))$, where $x_0 \in \mathbb{R}^2$ and $\delta r : [0, 2\pi] \to \mathbb{R}$ is a radial variation around a circle $\gamma_c(t)$ of radius r_0 centred at x_0. Then,

$$\gamma(t) = x_0 + \gamma_c(t) + \delta\gamma(t) \quad , \tag{7.1}$$

where $\gamma_c(t) = (r_0(t), \theta_c(t)) = (r_0, t)$, in Euclidean coordinates giving

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

We suppose that we are given an energy functional E defined on the space Γ of these curves. This energy functional depends on the image data as defined in the previous chapter and will be detailed below for this representation.

We recall here the procedure for constructing the single-object space without increasing its dimension. We consider an initial curve $\gamma \in \Gamma$; we then perform gradient descent to arrive at a local minimum of E, giving a second curve, $\tilde{\gamma} \in \Gamma$. The map $\tilde{\cdot} : \Gamma \to \Gamma$ takes every curve $\gamma \in \Gamma$ to a local minimum of the energy E. We define the space $\mathcal{C}$ of circles lying in the image domain, with radii in $[r_{\min}, r_{\max}]$, parameterized by arc-length. We consider now the single-object space $\Gamma_o = \tilde{\mathcal{C}}$ as a set of objects locally adapted to the data. Then despite allowing for potentially arbitrarily shaped star domains, the dimension of the single-object space remains small: an object is defined by a 'point', the center of a circle in the domain V and the radius of this circle, which is thus for one-dimensional object 'mark'.

To define the space of possible single objects Γ_o, we have to define the energy E. As mentioned in Introduction, we define it as the sum of a term related to the image and a term including prior information about object shape:

$$E(\gamma) = E_{\text{image}}(\gamma) + E_{\text{curve}}(\gamma) \quad . \tag{7.2}$$

The image energy term is defined as a weighted sum:

$$E_{image}(\gamma) \quad = \quad \lambda_g E_{grad}(\gamma) + \lambda_G E_{gauss}(\gamma) \quad .$$

where the terms are defined as follows:

$$E_{\text{grad}}(\gamma) = \int_{[0,2\pi]} dt\, n(\gamma(t)) \cdot \nabla I(\gamma(t)) \quad , \tag{7.3}$$

$$E_{\text{gauss}}(\gamma) = \int_{R(\gamma)} d^2\mathrm{x}\, (G(\mathrm{x}) - \bar{G}(\mathrm{x})) \quad , \tag{7.4}$$

recall that $n(\gamma(t))$ is the (unnormalized) outward normal to the curve; I is the image; $G(\mathrm{x}) = \frac{(I(\mathrm{x})-\mu)^2}{2\sigma^2}$ and $\bar{G}(\mathrm{x}) = \frac{(I(\mathrm{x})-\bar{\mu})^2}{2\bar{\sigma}^2}$; and $R(\gamma)$ is the region surrounded by the curve γ. The first term favours objects with high image gradients normal to their boundaries. The second one comes from a Gaussian model of the image, with parameters μ, σ, and $\bar{\mu}$, $\bar{\sigma}$, estimated using the object and background fragments of image.

Thus E_{image} is the same as in the previous model, whereas E_{curve} is now defined as the sum of two terms:

$$E_{\text{curve}}(\gamma) = E_{\text{smth}}(\gamma) + E_{\text{sh}}(\gamma) \quad . \tag{7.5}$$

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

$$E_{\text{curve}}(\gamma) = \frac{1}{2\pi} \iint_{[0,2\pi]} dt dt' \; F(t-t')\delta r(t)\delta r(t') \quad . \tag{7.6}$$

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

$$E_{\text{smth}}(\gamma) = \int_{[0,2\pi]} dt \, |\dot{\gamma}(t)|^2 \quad , \tag{7.7}$$

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, with the Fourier coefficients having zero mean except at zero frequency, which corresponds to a change in radius, and is absorbed in r_0. The energy is thus defined by the variance of each Fourier component, $g(k) = \frac{1}{4\pi\sigma(k)^2}$. 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:

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

where

$$\hat{\delta r}(k) = \frac{1}{2\pi} \int_{[0,2\pi]} dt \exp(-ikt)\delta r(t) \quad . \tag{7.8}$$

Figures 7.1 and 7.2 show examples of shapes obtained as a circle of radius r_0 with variations $\hat{\delta r}(k) = \sigma(k)$ at different frequencies k.

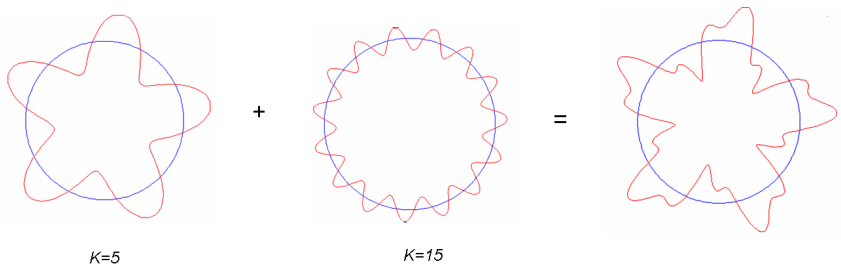

Figure 7.1: Shape obtained as a perturbation of a circle at frequencies $k = \{5, 15\}$, with $\hat{\delta r}(5) = 5$, and $\hat{\delta r}(15) = 2$, where $r_0 = 20$.

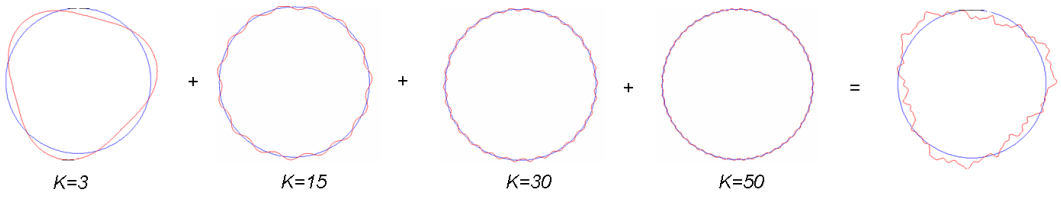

Figure 7.2: Shape obtained as a perturbation of a circle at frequencies $k = \{3, 15, 30, 50\}$, with $\hat{\delta r}(3) = 1.5$, $\hat{\delta r}(15) = 0.5$, $\hat{\delta r}(30) = 0.3$, $\hat{\delta r}(50) = 0.2$, where $r_0 = 20$.

Notice that for our parameterized set of curves, the curve derivative in Euclidean coordinates is given as follows:

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

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

$$|\dot{\gamma}(t)|^2 = \dot{\delta r}(t)^2 + (r_0 + \delta r(t))^2 \quad , \tag{7.9}$$

so that we can write:

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

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

$$\begin{aligned} E_{\text{curve}}(\gamma) \quad &= \quad 2\pi \sum_{k\in\mathbb{Z}} (k^2 + 1 + g(k))|\hat{\delta r}(k)|^2 \\ &\quad\quad +2\pi 2 r_0 \hat{\delta r}(0) + 2\pi r_0^2 \quad , \end{aligned}$$

taking into account that $\dot{x}_0^x(t) = \dot{x}_0^y(t) = 0$, defining a function $f(k) = k^2 + 1 + g(k)$, and dropping both $2\pi r_0^2$, which is simply an additive constant, and the linear term $2\pi 2 r_0 \hat{\delta r}(0)$, which serves only to change the mean of $\hat{\delta r}(0)$, which we define to be zero, we can write:

$$E_{\text{curve}}(\gamma(t)) = 2\pi \sum_{k\in\mathbb{Z}} f(k)|\hat{\delta r}(k)|^2 \quad . \tag{7.10}$$

Now using the inverse Fourier Transform of $f(k)$ and taking into account equation (7.8), equation (7.10) can be written as

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

Thus we obtain equation (7.6) where $F(t-t') = \sum_{k\in\mathbb{Z}} \exp ik(t-t') f(k)$.

The algorithm makes uses of the functional derivative of E:

$$\frac{\delta E}{\delta\gamma(t)} = \left(\frac{\partial E}{\partial x_0(\gamma)}, \frac{\delta E}{\delta\delta r(t)}\right) \quad , \tag{7.11}$$

the components of which are found as described below.

We define a clockwise traveling direction along the curve, so that an outward normal to the curve is given as follows:

$$\begin{aligned} n^x(\gamma(t)) &= -\dot{\gamma}^y(t), \\ n^y(\gamma(t)) &= \dot{\gamma}^x(t) \quad . \end{aligned} \tag{7.12}$$

Then, we define a vector field $v(\gamma)$ as a rotated image vector field $\nabla I(\gamma)$ as follows:

$$\begin{aligned} v^x(\gamma(t)) &= -\partial_y I(\gamma(t)) \\ v^y(\gamma(t)) &= \partial_x I(\gamma(t)) \end{aligned} \tag{7.13}$$

so that we can write:

$$E_{grad}(\gamma) = \int_{[0,2\pi]} dt\, n(\gamma(t)) \cdot \nabla I(\gamma(t)) = \int_{[0,2\pi]} dt\, \dot{\gamma}(t) \cdot v(\gamma(t)) \quad . \tag{7.14}$$

The components of the derivative $\frac{\partial E_{grad}}{\partial x_0(\gamma)}$ are then given by the following expressions:

$$\begin{aligned} \frac{\partial E_{grad}}{\partial x_0^x(\gamma)} &= \int_{[0,2\pi]} dt(\dot{\gamma}^x \partial_x v^x + \dot{\gamma}^y \partial_x v^y)(t) \\ &= \int_{[0,2\pi]} dt(-\dot{\gamma}^x \partial_x \partial_y I + \dot{\gamma}^y \partial_x \partial_x I)(t) \quad , \end{aligned} \tag{7.15}$$

and

$$\begin{aligned}\frac{\partial E_{grad}}{\partial x_0^y(\gamma)} &= \int_{[0,2\pi]} dt(\dot\gamma^x \partial_y v^x + \dot\gamma^y \partial_y v^y)(t) \\ &= \int_{[0,2\pi]} dt(-\dot\gamma^x \partial_y \partial_y I + \dot\gamma^y \partial_y \partial_x I)(t) \ . \end{aligned} \tag{7.16}$$

In the same way that we previously computed $\frac{\partial E_{grad}}{\partial x_0(\gamma)}$, we define now a function $h(\mathrm{x}) = \nabla \cdot U(\mathrm{x}) = (G - \bar G)(\mathrm{x})$ in order to present (7.4) using Green's theorem as follows:

$$\begin{aligned} E_{gauss}(\gamma) &= \int_{R(\gamma)} d^2\mathrm{x}\ (G(\mathrm{x}) - \bar G(\mathrm{x})) \\ &= \int_{R(\gamma)} d^2\mathrm{x}\ h(\mathrm{x}) && (7.17) \\ &= \int_{R(\gamma)} d^2\mathrm{x}\ \nabla \cdot U(\mathrm{x}) && (7.18) \\ &= \int_{\partial R(\gamma)} dt\ n(\gamma(t)) \cdot U(\gamma(t)) \ , && (7.19) \end{aligned}$$

i.e. it takes thus the same form as the image gradient term (7.14). Using a curve derivative we can re-write it as:

$$\int_{\gamma(t)} dt\ n(\gamma(t)) \cdot U(\gamma(t)) = \int_{\gamma(t)} dt\ \dot\gamma(t) \cdot v(\gamma(t)) \ ,$$

where

$$\begin{aligned} v^x(\gamma(t)) &= -U^y(\gamma(t)) \\ v^y(\gamma(t)) &= U^x(\gamma(t)) \ . \end{aligned} \tag{7.20}$$

The components of $\frac{\partial E_{gauss}}{\partial x_0(\gamma)}$ thus take the form:

$$\begin{aligned}\frac{\partial E_{gauss}}{\partial x_0^x(\gamma)} &= \int_{[0,2\pi]} dt \left(\dot\gamma^x(t) \partial_x v^x(\gamma(t)) + \dot\gamma^y(t) \partial_x v^y(\gamma(t))\right) \\ &= \int_{[0,2\pi]} dt \left(- \dot\gamma^x(t) \partial_x U^y(\gamma(t)) + \dot\gamma^y(t) \partial_x U^x(\gamma(t))\right) \ , \end{aligned} \tag{7.21}$$

and

$$\begin{aligned}\frac{\partial E_{gauss}}{\partial x_0^y(\gamma)} &= \int_{[0,2\pi]} dt \left(\dot\gamma^x(t) \partial_y v^x(\gamma(t)) + \dot\gamma^y(t) \partial_y v^y(\gamma(t))\right) \\ &= \int_{[0,2\pi]} dt \left(- \dot\gamma^x(t) \partial_y U^y(\gamma(t)) + \dot\gamma^y(t) \partial_y U^x(\gamma(t))\right) \ . \end{aligned} \tag{7.22}$$

We consider $h(\gamma) = \nabla \cdot U(\gamma)$, and assume then $\partial_x U^x(\gamma) = ph(\gamma)$ and $\partial_y U^y = qh(\gamma)$, where $p + q = 1$, so that we can write:

$$\begin{aligned} U^x &= p \int_0^x dx'\ h(x', y) \\ U^y &= q \int_0^y dy'\ h(x, y') \quad . \end{aligned}$$

The mixed derivatives of U thus take the following form:

$$\begin{aligned} \partial_y U^x &= p \frac{\partial}{\partial y} \int_0^x dx'\ h(x', y) \\ \partial_x U^y &= q \frac{\partial}{\partial x} \int_0^y dy'\ h(x, y') \quad . \end{aligned}$$

The energy derivative w.r.t. to x_0 can now be written as

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

with U a vector field satisfying $\nabla \cdot U = (G - \bar{G})$ and $i = \{x, y\}$.

The components of the energy functional derivative w.r.t $\delta r(t)$ are given in following way. For the prior curve term, we obtain:

$$\frac{\delta E_{curve}}{\delta \delta r(t)} = \frac{1}{\pi} \int_{[0,2\pi]} dt'\ F(t - t') \delta r(t') \quad . \tag{7.23}$$

The image gradient term, taking into account the expression for $\dot{\gamma}(t)$, gives

$$\begin{aligned} E_{grad}(\gamma) &= \int_{[0,2\pi]} dt\ n(\gamma(t)) \cdot \nabla I(\gamma(t)) = \int_{[0,2\pi]} dt\ \dot{\gamma}(t) \cdot v(t) \\ &= \int_{[0,2\pi]} dt [(\dot{\delta r}(t) \cos(t) - (r_0 + \delta r(t)) \sin(t)) v^x(t) \\ &\quad + (\dot{\delta r}(t) \sin(t) + (r_0 + \delta r(t)) \cos(t)) v^y(t)] \quad ; \end{aligned} \tag{7.24}$$

then its derivative can be written as:

$$\begin{aligned} \frac{\delta E_{grad}}{\delta \delta r(t)} &= (r_0 + \delta r(t))(\partial_x v^y(\delta r(t)) - \partial_y v^x(\delta r(t))) \\ &= (r_0 + \delta r(t)) \left[\partial_x^2 I + \partial_y^2 I\right](\gamma(t)) \quad . \end{aligned} \tag{7.25}$$

For the Gaussian term, taking into account, that

$$E_{gauss}(\gamma) \quad = \quad \lambda_G \int_{R(\gamma)} d^2\mathrm{x}\,(G(\mathrm{x}) - \bar{G}(\mathrm{x})) = \int_{\gamma(t)} dt\,\dot{\gamma}(t) \cdot v(\gamma(t)) \quad ,$$

the derivative is computed analogously

$$\begin{aligned} \frac{\delta E_{gauss}}{\delta\delta r(t)} \quad &= \quad (r_0 + \delta r(t))(\partial_x v^y(\delta r(t)) - \partial_y v^x(\delta r(t))) \\ &= (r_0 + \delta r(t))(\partial_x U^x(\delta r(t)) + \partial_y U^y(\delta r(t))) \\ &= (r_0 + \delta r(t))(G - \bar{G})(\gamma(t)) \quad . \end{aligned} \tag{7.26}$$

Taking now the sum of the components (7.15), (7.16), (7.21), and (7.22), the total energy derivative w.r.t. to $\delta r(t)$ takes form:

$$\begin{aligned} \frac{\delta E}{\delta\delta r(t)} \quad &= \quad \lambda_g(r_0 + \delta r(t))\nabla^2 I(\gamma(t)) \\ &\quad +\lambda_G(r_0 + \delta r(t))(G - \bar{G})(\gamma(t)) \\ &\quad +\frac{1}{\pi}\int_{[0,2\pi]} dt' F(t - t')\delta r(t') \quad . \end{aligned} \tag{7.27}$$

For the computational details, the reader can refer to the appendix B.2.

Figures 7.3 and 7.4 show gradient descent experiments using the energy E. For the first experiment, the function $g(k)$ is defined so as to discourage all frequencies: $g(k) = 2$, $k = 0, 2 \ldots, N/2$, while for $k = 1$, $g(k)$ is fixed to the maximum machine value, since $\hat{\delta r}(1)$ acts as a translation. The model thus favours delineation by a curve close to a circle with radius $r_0 = 75$. For the second experiment, the frequency $k = 4$ was relaxed ($g(4) = 0.0002$).

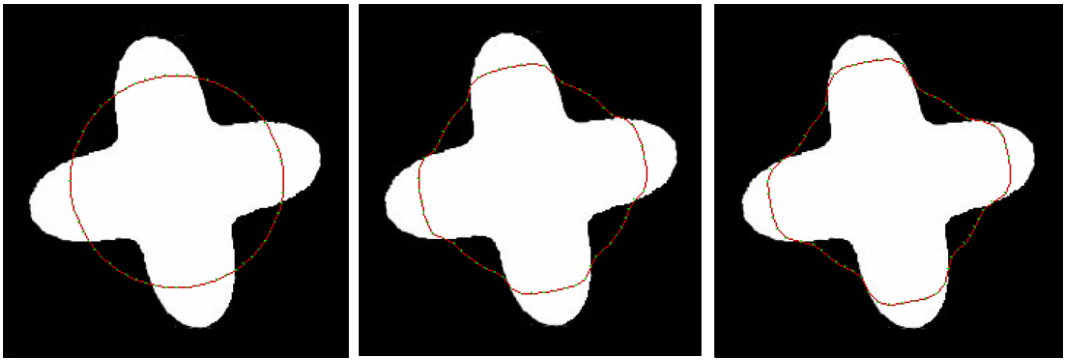

Figure 7.3: Synthetic image example: gradient descent driven by E, where $g(k)$ is defined so as to discourage variations around a circle of radius r_0.

Figure 7.4: Synthetic image example: gradient descent driven by E, where $g(k)$ is defined so as to discourage the variations around a circle of radius r_0, except that of frequency $k = 4$.

7.3 Experimental results

In this section we present several results obtained using the MPP model for arbitrarily-shaped multiple object extraction with strong object shape prior incorporated. The experiments were performed on one band of very high resolution colour infra-red aerial images that represent scenes composed of visible top part of tree crowns. The images show natural forest zones, the trees thus being of different species, age and sizes; many of them are grouped into twos, threes or more. We show, first, a result obtained on a fragment of an image with distinct as well as overlapping crowns, and visible as well as partly cluttered crowns, the trees being of different species and thus of diverse shape *e.g.* spiky spruces and smoother birches and spruces again. It is a representative example of an image scene, where the differences in the accuracy of delineation and the number of detected objects between the two models can be simply viewed and compared. We then show experimental results on another part of an image where the trees are of remarkably different sizes; configurations at different iterations of the birth-and-death process are shown. Finally, we present the results of experiments on complete images (900x900 pixels) with different scene complexities. For Model II, the parameters of the annealing scheme were fixed for all the experiments: the time-discretisation step update $\Delta\delta = 0.99$ and the inverse temperature update $\Delta\beta = \frac{1}{0.993}$. The parameters of the image model, *i.e.* $\sigma, \mu, \bar{\sigma}, \bar{\mu}$, are those learned using examples of object and background for the tests of Model I. The radius range depends on the size of object sought relative to the image resolution, and thus can be easily calibrated. The weighting parameters are learned experimentally.

Visible tree crowns seen from vertically above have roughly circular shape with perturbations corresponding to the branches or leaves. We define thus the function $g(k)$ so as to discourage low frequencies, and in particular $k = 2$, in order to avoid the extraction of two overlapping trees as a single elongated object, as well as to favour roughly circular objects with small perturbations corresponding to branches or leaves ($g(0) = 0.04$, $g(2) = 2$, $g(3) = 0.96$; $g(k) = 1.2$ for $k = 4, \ldots, 12$; $g(k) = 0.0015$, when $k = 13, \ldots, N/2$).

We might think that it is easy to estimate the function $g(k)$ because of the choice of a Gaussian model and the available set of tree crown shapes studied in Part I of the thesis. However, most of the shapes from that set are not "star-shapes", which is because they were extracted so as to preserve the maximum of details. A solution of this problem, which can be seen as a future work, would consist either in simplifying the shapes so as to make them star-domains; or in further studying them in order to develop another more appropriate model for tree crown shape modeling.

As mentioned above, we demonstrate first the results of object extraction on a fragment of an image (cf Fig. 7.5), that was also used to test the model without the strong shape prior. The figure shows that the strong shape information allows us to extract nearby and even overlapping objects without paying the price of greatly simplifying their geometry.

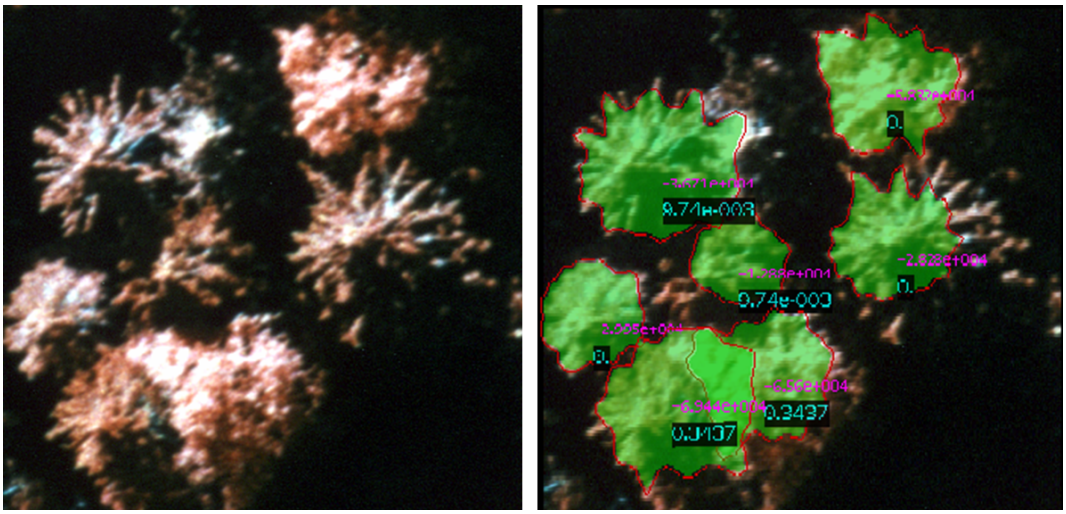

Figure 7.5: Left: original CIR image of tree crowns, © CBA. Right: final configuration obtained. The numbers show the interaction term value H_2 (black background) and the data term value H_1.

Figure 7.6 shows results obtained on an image fragment, where the crowns are of comparable sizes except one (pine) overlapping with another crown in the middle top part of the image. The pine crown is approximately twice as big as other contours in the image. Since a shape prior keeps the curve as close as possible to a circle of an average radius, the curve can not thus evolve so as to delineate the whole pine crown, which is finally detected as two slightly overlapping objects adjusted to the crown boundary. We note also, that during the iterations, the crowns of bigger size are detected first. They have lower energy, essentially because of the Gaussian term. Then, as the temperature decreases, the smaller objects 'survive'.

Finally, figures 7.7 and 7.8 demonstrate the results of the experiments on other images of 900x900 pixels using the same model parameters as for the previous experiments, except that the radius range $[r_{\min}, r_{\max}]$ is wider and is defined w.r.t. to these two images. The images represent complex scenes that contain mostly overlapping visible as well as occluded tree crowns of various sizes. The results show the overlapped crowns of different species and sizes detected as distinct objects quite accurately, whereas delineation becomes a very challenging task when the crowns are of the same species, being probably of approximately the same age and grouped together, *e.g.* a group of birch crowns in the middle right part of the image in Fig. 7.7. Here the role of the strong shape prior becomes crucial, and even

more so for the case of aspen trees, where the crowns are composed of random light and dark disjoint regions, cf left upper part of the original image in Fig. 7.7, and especially the bottom of the image in Fig. 7.8. In this case, it is impossible to delineate the crowns even manually, at least for a non-expert. These groups of trees were recovered as distinct overlapping objects, which were placed at the top parts of the crowns, where the vegetation is more dense, reflecting thus more infrared light. As in the previous experiments, we can see the tendency of to add smaller crowns to the current configuration as the temperature decreases. The final configurations obtained shown in the figures miss some small parts of crowns of trees. These could be young and thus have a very small visible part of the crown, or could be shadowed by the neighbouring trees, or lower branches. These tree fragments are not detected as whole distinct objects, they have a size and a shape far from the object shape considered and controlled by the shape prior term included. The computation time of the C++ programme for these two images was 85mn on a 2.16GHz processor.

7.4 Summary

In the previous chapter, we presented the extension to arbitrarily shaped objects of the marked point process framework used previously for the extraction of objects with simple geometries from images. In this chapter, 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 data and strong prior shape information about the objects to be segmented. This allows us to deal with configurations of overlapping objects with complex shapes. 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.

7.4.1 Comparison of models using different shape priors

In order to demonstrate the evolution of marked point process models taking now into account the models developed in our work, we present the results of object extraction obtained using three different models, cf Fig. 7.9. The top right image in the figure shows the configuration obtained using simply-shaped objects. For this experiment, the 2D model for tree crown extraction with objects represented by ellipses was used as described in [Perrin, 2006]. The optimisation procedure was, however, performed by sampling from a multiple birth-and death process. The data energy term in this model is defined as a function of the Bhattacharya distance between the grey level distributions of the set of pixels belonging to the object interior and those of the object border of some fixed size, whereas the interaction energy term controls the overlapping of objects. The bottom left image of the figure shows the

configuration obtained using the representation and energy described in the previous chapter, with no strong shape information included, *i.e.* with $E_{\text{sh}} \equiv 0$, or equivalently, $g \equiv 0$. The bottom right image shows the result obtained using the energy described in this chapter, *i.e.* with strong prior shape information included.

An MPP model using simple objects allows the rapid extraction of the objects from an image (4 mn for the first experiment and approximately 15 mn for the second and third experiments), 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 lower resolution images. The second approach, an MPP model for the extraction of arbitrarily-shaped objects with only weak shape information, is geometrically far more accurate, while not increasing the computational complexity unduly. The limitations of this approach, however, are that because it uses only 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 deals with both these limitations. For example, in the bottom left corner of the bottom left image in figure 7.9, 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 chapter.

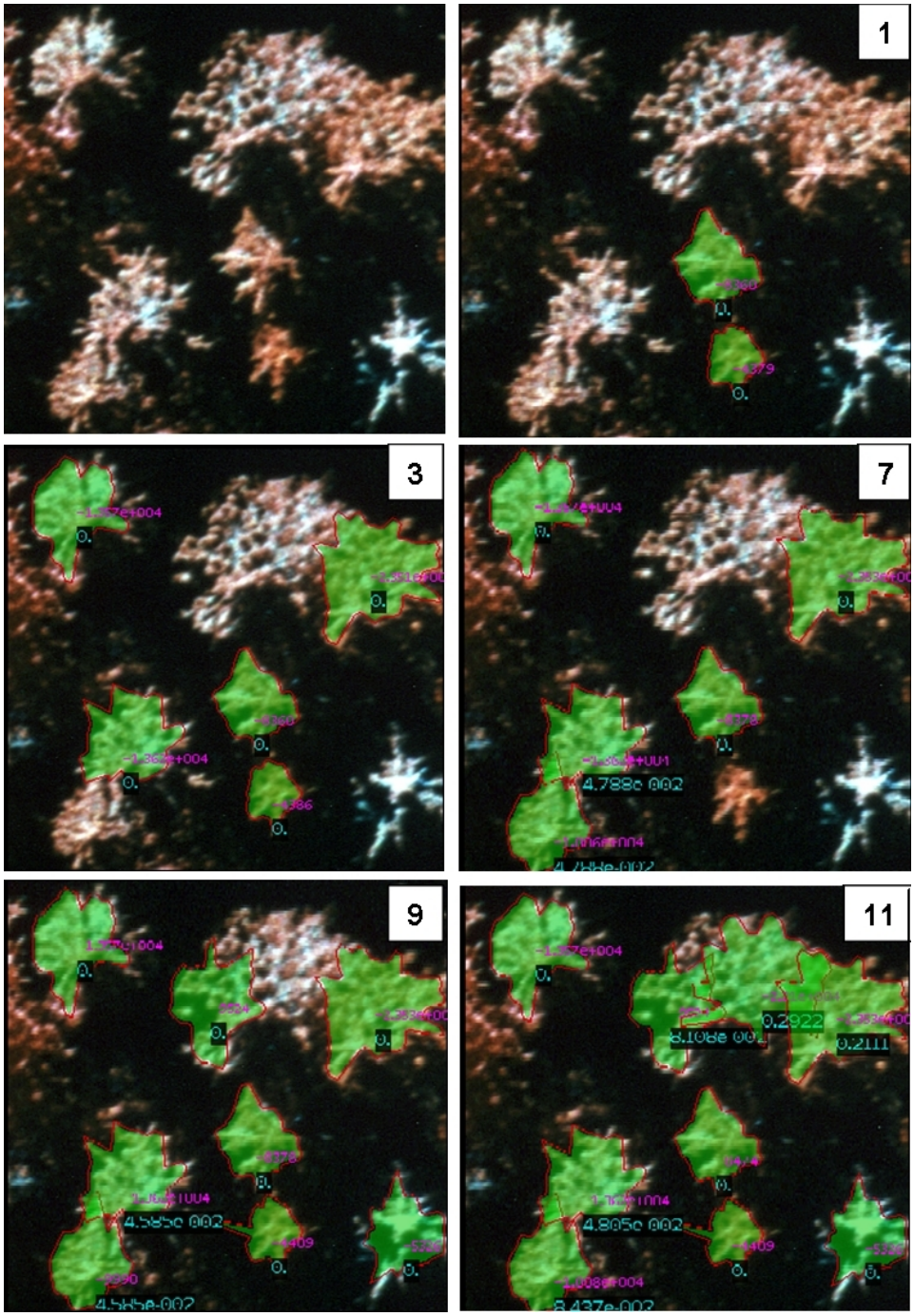

Figure 7.6: Top, left: original CIR image of tree crowns, © CBA. Bottom, right: final configuration obtained. The intermediate images show the configuration evolution during the birth-and-death process at iterations numbered in the top right corner of the images. The numbers on the crowns show the interaction term value H_2 (black background) and the data term value H_1.

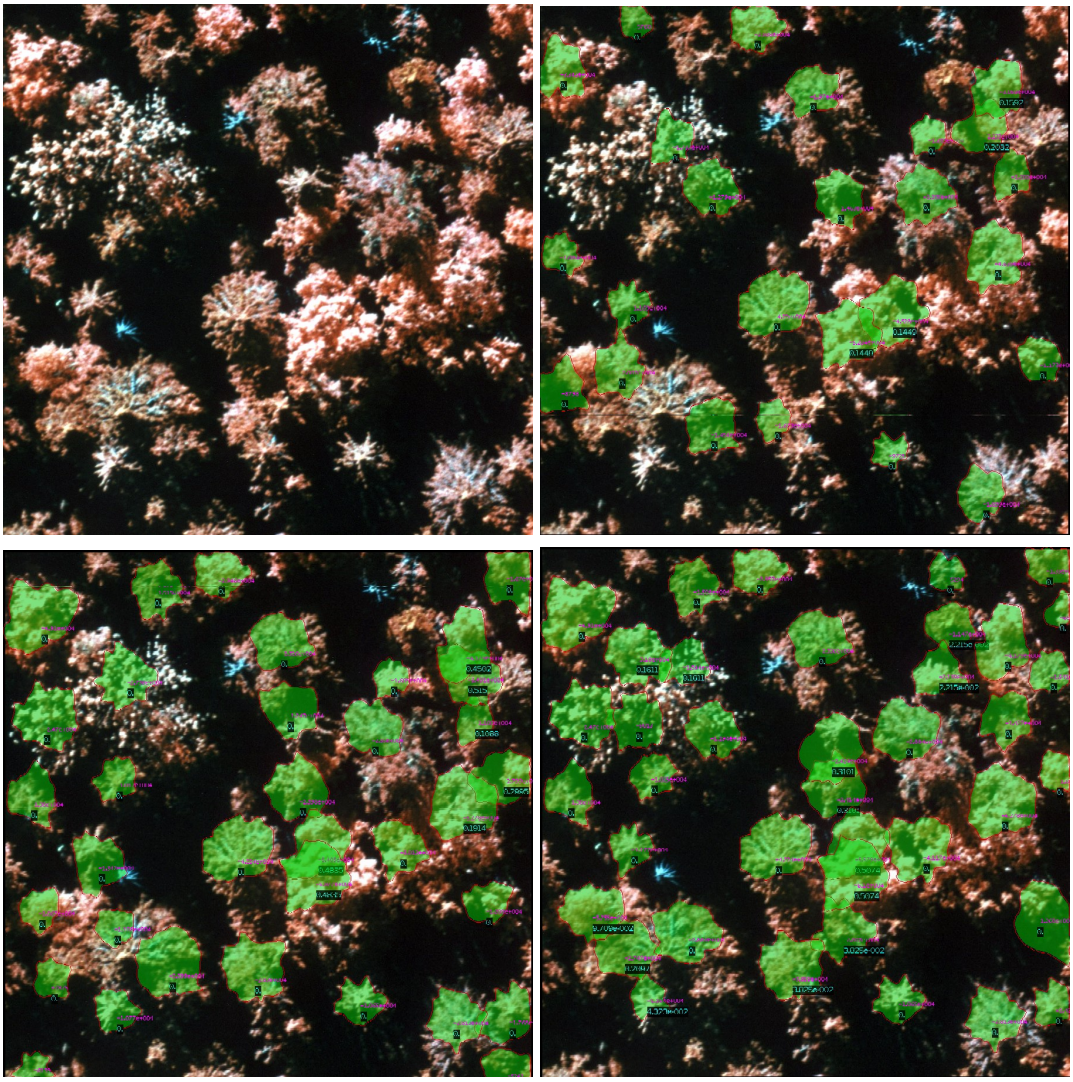

Figure 7.7: Top, left: original CIR image of tree crowns, © CBA. Bottom, right: final configuration obtained. The intermediate images show the configuration evolution during the birth-and-death process. The numbers on the crowns show the interaction term value H_2 (black background) and the data term value H_1.

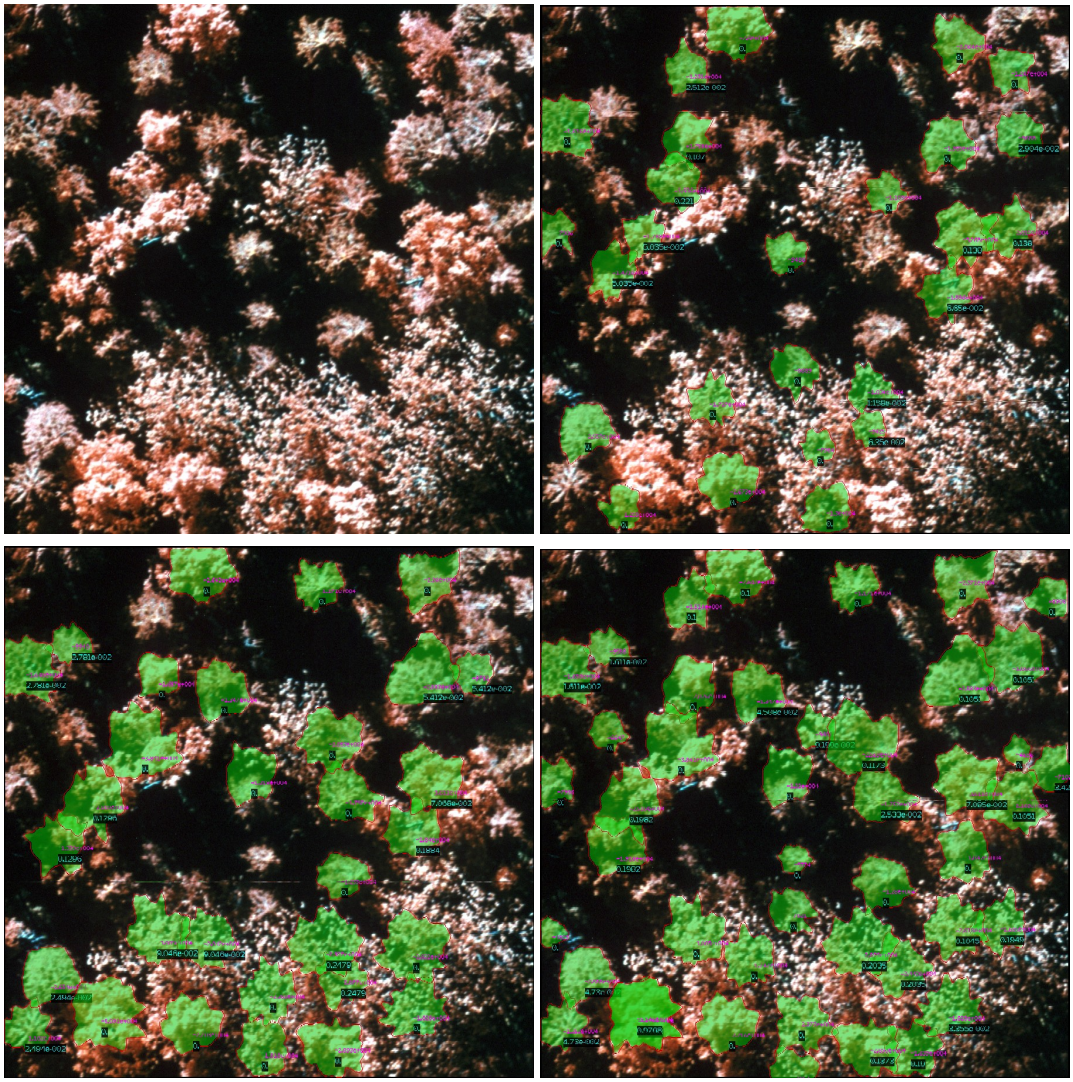

Figure 7.8: Top, left: original CIR image of tree crowns, © CBA. Bottom, right: final configuration obtained. The intermediate images show the configuration evolution during the birth-and-death process. The numbers on the crowns show the interaction term value H_2 (black background) and the data term value H_1.

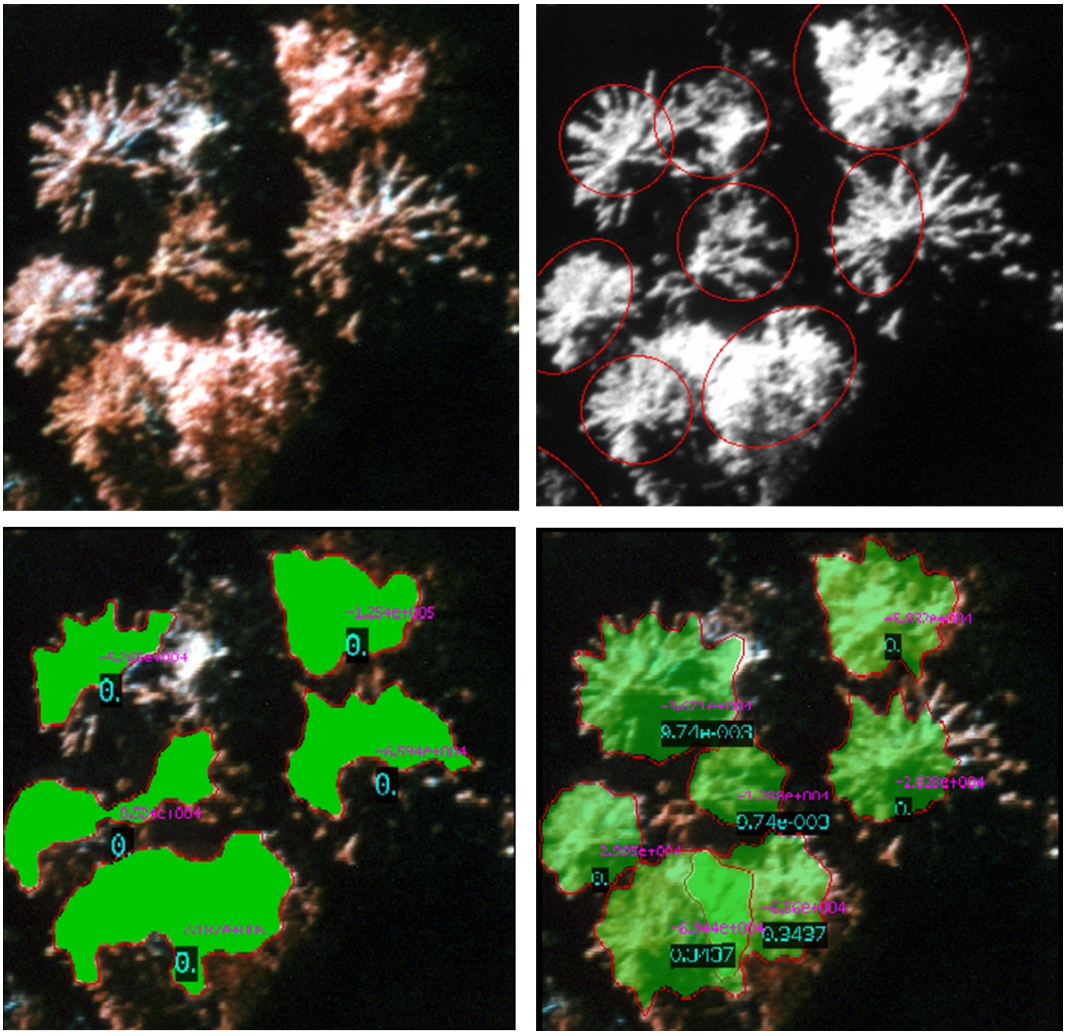

Figure 7.9: Top, left: original CIR image of tree crowns, © CBA. Top, right: final configuration using ellipse-shaped objects. Bottom, left: final configuration using arbitrarily-shaped objects obtained using an MPP without shape prior. Bottom, right: final configuration of arbitrarily-shaped objects 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.

Conclusion - Perspectives

In part I of this manuscript, we presented a study based on shape information (cf section 2.2) completed by radiometric and textural characteristics of tree crowns for their classification into species (cf section 3.4), the trees being of the four most prevalent species in Swedish forest. The classification is based on crowns delineated from colour infra-red 3 cm resolution aerial images of forest zones.

In part II, we presented a first extension of the marked point process framework for the automatic extraction from images of multiple objects with simple geometric shape to the extraction of objects with arbitrary, complex shape. Then we proposed a method for incorporating strong prior information about the shape of the objects that we wish to extract (cf section 7.4).

Referring to the topics addressed during this PhD thesis, the perspectives, conceivable as longer term goals, offered by this work are described below.

As seen in part I, tree crowns vary greatly in shape between classes as well as within each class due to the fact that natural forests are composed of trees of diverse species and age. Photographs of such forest zones represent thus very non-homogeneous scenes. While applying our model to such images, the model is limited by its general shape prior. Concerning this particular application, for more accurate delineation of tree crowns necessary, for instance, for tree crown classification, we propose to further extend the model to the extraction of arbitrarily-shaped objects by considering several classes. In this case, the space of single-objects consists of objects locally adapted to the data with different shape priors included. An object has thus an additional mark, a label l^{ω_i} corresponding to the class defining the shape prior. Therefore, every object now defined by $\omega_i = (x_0^{\omega_i}, r_0^{\omega_i}, l^{\omega_i})$ is obtained via gradient descent with the energy H_1 from a circle with centre $x_0^{\omega_i}$ and radius $r_0^{\omega_i}$, where H_1 includes the shape prior information to which l^{ω_i} corresponds. The energy $H(\omega)$ can be then written:

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

where ω is a configuration of objects ω_i.

Here, we present a first result using such a model on a fragment of a one band CIR image of two tree crowns. We have chosen for this experiment two examples of birch and spruce tree crowns, cf figures 8.13 and 8.14, which show the evolution of a curve under gradient descent with energy H_1 with two different shape priors: a 'birch' prior, with $g(k)$ chosen to encourage curves with perturbations of small amplitude corresponding to the leaves, and a 'spruce' prior with $g(k)$ favouring only low frequency perturbations representing the branches. The labels are then $l = \{$'birch', 'spruce'$\}$.

Figure 8.15 shows extraction of two crowns with two shape priors included. In both cases, using the right prior leads to a better delineation of the crown.

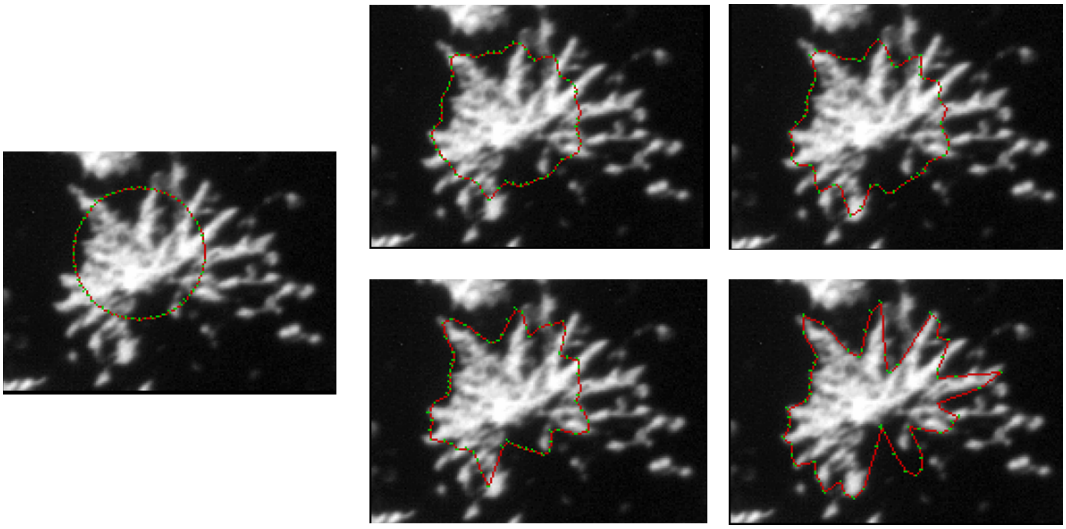

Figure 8.10: Evolution of a curve under gradient descent of energy H_1 with the 'birch' shape prior (top) and the 'spruce' shape prior (bottom). The image represents a spruce tree crown.

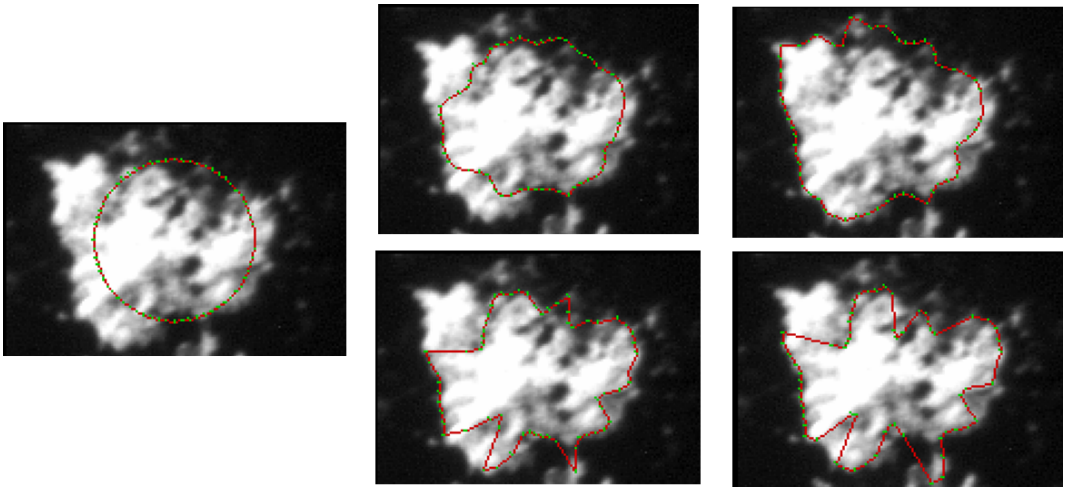

Figure 8.11: Evolution of a curve under gradient descent with energy H_1 with the 'birch' shape prior (top) and the 'spruce' shape prior (bottom). The image represents a birch tree crown.

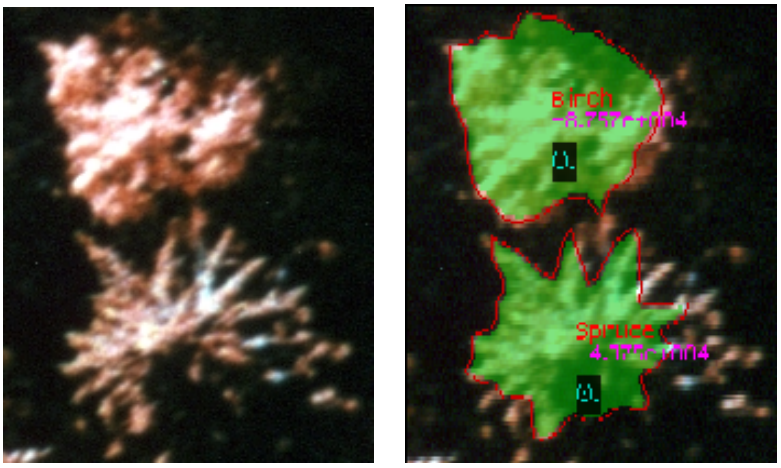

Figure 8.12: Left: a fragment of an original CIR image of tree crowns, © CBA. Right: tree crowns extracted with 'birch' or'spruce' shape priors included. The numbers on the crowns show the interaction term value H_2 (black background) and the data term value H_1.

The data energy term $H_1(\omega_i) = E(\tilde{\omega}_i)$ is written in the form of an active contour, which allowed us to include more specific shape prior information into the single-object model. We could thus think to add another type of information, for example, radiometric or textural, typical to one or another class, so that a single object can be presented as $\omega_i = (x_0^{\omega_i}, r_0^{\omega_i}, l_{sh}^{\omega_i}, l_{txt}^{\omega_i})$, where $l_{sh}^{\omega_i}$ is a label for a class of shapes, and $l_{txt}^{\omega_i}$ - for a class depending on texture. This type of model could be seen as a joint tree extraction and classification algorithm.

The models developed include numerous parameters which were calibrated (radius range for circle initialization), estimated from object and background examples (parameters arising from the Gaussian image model), or learned from experiments (weighting parameters). Another possible continuation of this work consists in estimating these parameters. Classical classification methods such as k-means allow the estimation of, for instance, the parameters related to the data. In [Chatelain *et al.*, 2009], a method for estimating the temperature T of the annealing scheme and the parameter weighting the data and interaction energy terms is proposed in the case of a simple marked point process using discs. Developing a fully automatic approach is nevertheless an open and challenging problem.

Dans la première partie du manuscrit, nous présentons une étude fondée sur l'information des formes (cf sous-chapitre 2.2), et complétée par les caractéristiques radiométrique et de texture des couronnes des arbres pour classifier, selon leur espèce, des arbres appartenant aux quatres espèces les plus représentées dans les forêts suédoises (cf sous-chapitre 3.4). La classification est fondée sur les couronnes délinées d'images aériennes infra-rouge couleur de zones forestières à 3 cm de résolution frounies pas l'Uiversité des sciences agricoles de Suède.

Dans la seconde partie de la thèse, nous introduison tout d'abord une extension du modèle de processus ponctuels marqués pour l'extraction automatique de multiples objets de formes géométriques simples à des objets de formes arbitrairement complexes à partir d'une image. Puis, nous proposons une méthode pour inclure l'information a priori sur la forme des objets à extraire (cf sous-chapitre 7.4).

Se référant aux thèmes abordés au cours de cette thèse, différentes perspectives peuvent être envisagées.

Comme nous l'avons vu dans la partie I, la forme des couronnes des arbres varie grandement selon les classes; et même selon les arbres de chaque classe. Cela est dû au fait que les forêts naturelles sont composées d'arbres d'espèces et d'âges différents. Les photographies de telles forêts représentent donc des scènes très hétérogènes. L'application à de telles images du modèle présenté est donc limitée par la généralité du critère de forme a priori. Pour cette raison, nous proposons ici d'étendre le modèle d'extraction de formes arbitraires en considérant maitenant plusieurs classes. Ainsi, l'espace d'objets simples est constitué d'objets adaptés localement aux données avec des informations a priori différentes. L'objet a donc une marque supplémentaire, une étiquette l^{ω_i} correspondant à la classe définissant l'a priori sur la forme incorporée. Alors, chaque objet, maintenant défini par $\omega_i = (x_0^{\omega_i}, r_0^{\omega_i}, l^{\omega_i})$, est obtenu via une descente de gradient de l'énergie H_1 à partir d'un cercle de centre $x_0^{\omega_i}$ et de rayon $r_0^{\omega_i}$, où H_1 contient l'information a priori sur la forme à laquelle correspond l^{ω_i}. Une configuration ω est composée d'un ensemble d'objets ω_i. L'énergie $H(\omega)$ peut donc s'écrire comme :

$$H(\omega) = c_0 \sum_i \sum_l H_1^l(\omega_i)\, \delta(l^{\omega_i}, l) + \sum_{i \neq j} H_2(\omega_i, \omega_j) \quad .$$

Voici ci-dessous un premier résultat utilisant un tel modèle sur deux couronnes d'arbres dans un fragment d'image infra-rouge couleur mono canal. Les figures 8.13 et 8.14 montrent l'évolution de la courbe selon le gradient d'énergie H_1 contenant une information a priori sur la forme $g(k)$ du type 'bouleau' qui favorise des perturbations de petite amplitude correspondant aux feuilles, et une information a priori du type 'épicéa' dans laquelle l'a priori sur la forme favorise uniquement les perturbations de petites fréquences représentant les branches. Les formes sont étiquetées : $l = \{$'bouleau', 'épicé'$\}$.

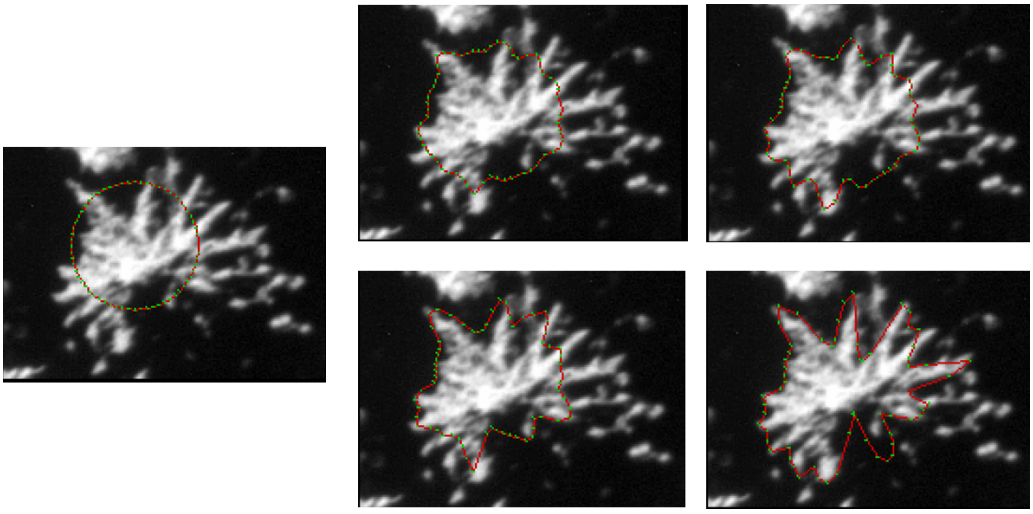

Figure 8.13: Evolution de la courbe par la descente de gradient de l'énergie H_1 avec l'a priori de forme du type 'bouleau' (en haut) et du type 'épicéa' (en bas). L'image représente une couronne d'épicéa. © CBA.

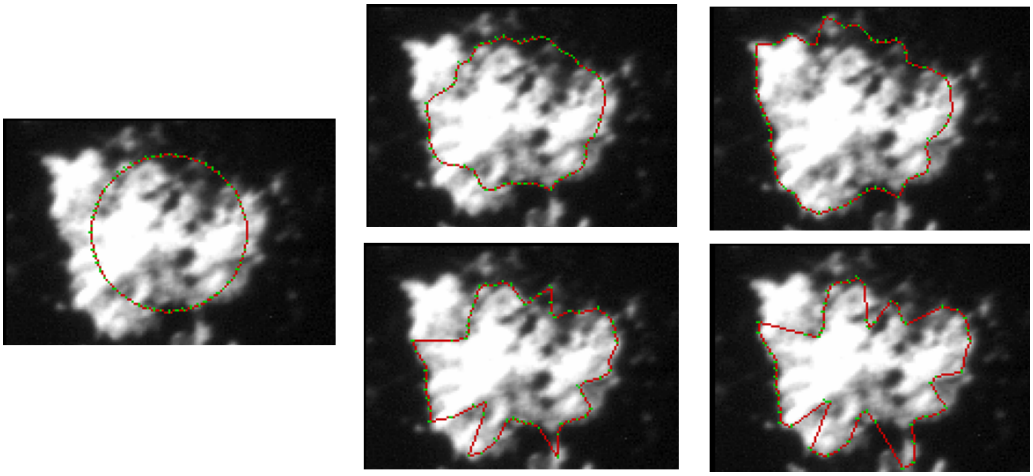

Figure 8.14: Evolution de la courbe par la descente de gradient de l'énergie H_1 avec l'a priori de forme du type 'bouleau' (en haut) et du type 'épicéa' (en bas). L'image représente une couronne de bouleau. © CBA.

La figure 8.15 montre l'extraction de deux couronnes d'arbres en utilisant deux informations a priori différentes. Dans les deux cas, l'utilisation de l'information a priori correspondant à l'espèce d'arbres conduit à une meilleure extraction de la forme.

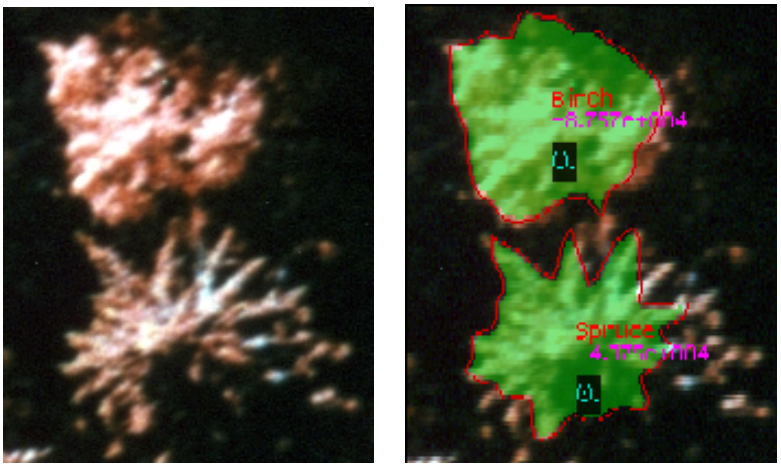

Figure 8.15: Gauche: un fragment de l'image originale IRC représentant des couronnes d'arbres, © CBA. Droite: couronnes extraites avec les deux a priori différents de types 'bouleau' et 'épicéa'. Les chiffres sur les couronnes montrent les valeurs des termes d'interaction H_2 (fond noir) et d'attache aux donnés H_1.

Le terme d'énergie d'attache aux données $H_1(\omega_i) = E(\tilde{\omega}_i)$ est écrit sous la forme d'un contour actif, ce qui permet d'inclure plus d'information a priori sur la forme dans le modèle d'objet individuel. Nous pourrions appliquer cette même méthode aux autres types d'attributs liés à l'image : la radiométrie, la texture caractéristique d'une classe ou l'autre, de sorte qu'un seul objet peut être présenté par $\omega_i = (x_0^{\omega_i}, r_0^{\omega_i}, l_{sh}^{\omega_i}, l_{txt}^{\omega_i})$, où $l_{sh}^{\omega_i}$ est une étiquette pour la classe de formes, et $l_{txt}^{\omega_i}$ une étiquette pour la classe de textures. Ce type de modèle peut-être considéré à la fois comme une technique d'extraction et de classification.

Les modèles développés utilisent de nombreux paramètres qui ont été calibrés (variation de la taille du rayon à l'initialisation des cercles), estimés à partir des objets et du fond de l'image (paramètres provenants du modèle supposé gaussien), ou empiriquement fixés au fur et à mesure des expériences (paramètres de pondération). Une suite possible de ce travail serait donc l'automatisation de l'estimation des paramètres. Les méthodes classiques de classification, comme la méthode des k-moyennes, permettent, par exemple, l'estimation des paramètres liés aux données. Une méthode d'estimation de la température dans le schéma du recuit simmulé, dans le cas de PPM simple utilisant des disques, est proposé dans [Chatelain *et al.*, 2009]. Néanmoins, le développement d'une approche complètement automatique reste un problème ouvert et un enjeu déterminant pour permettre une utilisation de ce modèle par un plus large public.

Appendix, Bibliography, and Résumé

Appendix A

Theory review

In this chapter we recall some theoretical issues that were not included in the main text.

A.1 Support Vector Machines

SVMs are a set of supervised learning methods for classification and regression. They work by finding the maximum separation (margin) between classes using a set of observations called the training data. SVMs are also called maximum margin classifiers since they minimize the empirical classification error and maximize the geometric margin simultaneously [Vapnik, 1998].

A.1.1 Linear SVM

Separable case

Given training data (x_i, y_i), $i = 1, ..., N$, where $x_i \in R^m$ and $y_i \in \{-1, 1\}$ denoting the class to which x_i belongs, SVM looks for the Optimal Separating Hyperplane that maximizes the distance between the closest training samples of the two classes, also called the 'margin'.

The separating hyperplane is defined as follows $w \cdot x + b = 0$. The vector w is a vector normal to the hyperplane and b is the offset parameter allowing us to increase the margin. So the classifier is given by $f : x \in R^m \longmapsto sign(w \cdot x + b) \in \{-1, 1\}$, i.e. all the training data satisfy the following constraints:

$$\begin{cases} w \cdot x_i + b \geq 0 \text{ if } y_i = +1 \\ w \cdot x_i + b \leq 0 \text{ if } y_i = -1 \end{cases}$$

These constraints can be described by a set of inequalities:

$$y_i(w \cdot x_i + b) \geq 1 \ , \ \forall i \tag{A.1}$$

To find the hyperplane that gives the maximum margin $2/\|w\|$ (for details see [Burgers, 1998]), we minimize $\|w\|^2$, subject to the constraints (A.1). This leads to the following quadratic optimization problem:

$$\min_{(w,b)} \frac{\|w\|^2}{2}$$

$$\text{subject to : } y_i(w \cdot x_i + b) \geq 1, \ \forall i.$$

By introducing the Lagrange multipliers $\lambda_i, i = 1, .., N$, one for each inequality constraint (A.1), we construct the dual problem [Burgers, 1998]:

$$\max_{\lambda} \left(L(\lambda) = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j y_i y_j x_i \cdot x_j \right),$$

$$\text{subject to } \sum_{i=1}^{N} \lambda_i y_i = 0 \ , \ 0 \leq \lambda_i \ , \ \forall i,$$

which is a convex quadratic optimization problem subject to linear constraints. Thus, the solution is given by $w = \sum_{i=1}^{N} \lambda_i y_i x_i$ and $b = y_i - w \cdot x_i$, for $i : \lambda_i \neq 0$. The classification function becomes:

$$f(x) = \sum_{i=1}^{N} \lambda_i y_i x_i \cdot x + b$$

Nonseparable case

For nonlinearly separable data, slack variables ξ_i and a regularization parameter C are introduced to deal with misclassified samples, i.e. to relax the constraints (see Fig. A.1). By introducing them into the optimization problem, we obtain:

$$\min_{(w,b)} \frac{\|w\|^2}{2} + C \sum_{i=1}^{N} \xi_i,$$

$$\text{subject to : } y_i(w \cdot x_i + b) \geq 1 - \xi_i \ , \ \xi_i \geq 0, \ \forall i.$$

Using the Lagrange multipliers, this quadratic problem becomes:

$$\max_{\lambda} \left(L(\lambda) = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j y_i y_j x_i \cdot x_j \right),$$

$$\text{subject to } \sum_{i=1}^{N} \lambda_i y_i = 0 \ , \ 0 \leq \lambda_i \leq C \ , \ \forall i.$$

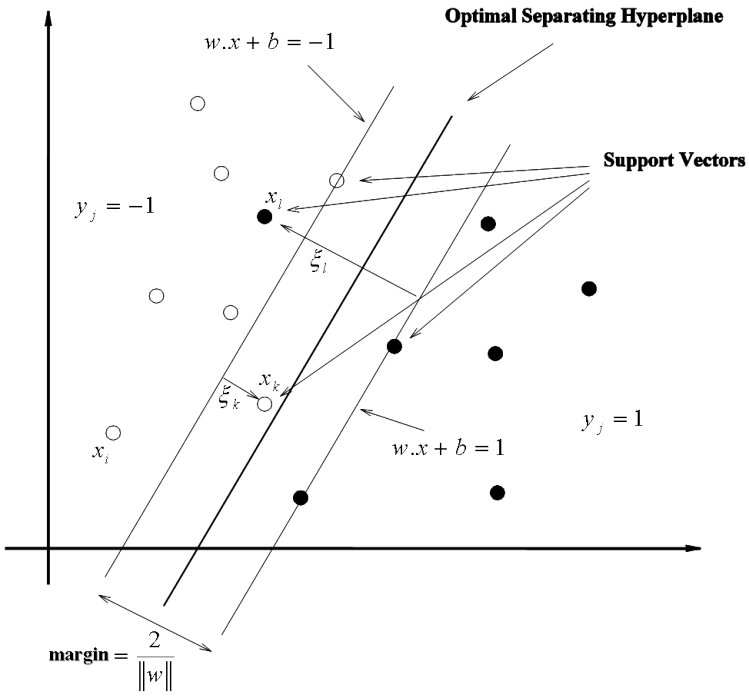

Figure A.1: SVM classifier.

A.1.2 Nonlinear SVM

For cases where the decision function is not a linear function of the data, a non-linear classifier can be created by applying the kernel trick (originally proposed by Aizerman) to maximum-margin hyperplanes (cf [Boser *et al.*, 1992]). The resulting algorithm is formally similar, except that every dot product $x_i \cdot x_j$ is replaced by a non-linear kernel function $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$, with $\Phi : R^m \longmapsto H$, H being a Euclidean space generally of a higher dimension. This allows the algorithm to fit the maximum-margin linear hyperplane in the space H.

The classification function then becomes:

$$f(x) = \sum_{i=1}^{N} \lambda_i y_i K(x_i, x) + b.$$

Note that there exists a mapping Φ if $K(x_i, x_j)$ satisfies the Mercer's condition [Vapnik, 1995].

Appendix B

Details of calculations

In this appendix, we detail the calculation of the variational derivatives of the energy terms used to perform gradient descent to define the single-object space for the marked point process models developed and presented in the second part of this thesis .

B.1 Model I

To compute the variational derivative of the image energy term $E_{image}(\gamma)$ defined by equation 6.2, we need to vary both of the terms it composing. It was shown, using Green's theorem, that the Gaussian term $E_{gauss}(\gamma)$ can be written in the same way as the image gradient term $E_{grad}(\gamma)$. Let us give the computation details for the image gradient term; the variation of the Gaussian term is computed analogously.

$$E_{grad}(\gamma) = \int_{[0,2\pi]} dt\, n(\gamma(t)) \cdot \nabla I(\gamma(t)) = \int_{[0,2\pi]} dt\, \dot{\gamma}(t) \cdot v(\gamma(t)) \quad , \tag{B.1}$$

where

$$\begin{aligned} v^x(\gamma(t)) &= -\partial_y I(\gamma(t)), \\ v^y(\gamma(t)) &= \partial_x I(\gamma(t)). \end{aligned}$$

Variation of (B.1) is given by

$$\begin{aligned} \delta_{\gamma(t)} E_{grad}(\gamma) &= E_{grad}(\gamma + \delta\gamma) - E(\gamma) \\ &= \int_{[0,2\pi]} dt\, (\dot{\gamma}(t) + \delta\dot{\gamma}(t)) v(\gamma(t)) + \int_{[0,2\pi]} dt\, \dot{\gamma}(t) \delta_{\gamma(t)} v(\gamma(t)) - \int_{[0,2\pi]} dt\, \dot{\gamma}(t) \cdot v(\gamma(t)) \\ &= \int_{[0,2\pi]} dt\, \delta\dot{\gamma}(t) v(\gamma(t)) + \int_{[0,2\pi]} dt\, \dot{\gamma}(t) \delta_{\gamma(t)} v(\gamma(t)) \quad . \end{aligned} \tag{B.2}$$

Integration by parts of the x component of the first term of (B.2) gives:

$$\begin{aligned}\int_{[0,2\pi]} dt\, \delta\dot{\gamma}^x(t) v^x(\gamma(t)) &= \frac{d}{dt} v^x(\gamma(t)) \delta\gamma^x(t)\big|_0^{2\pi} - \int_{[0,2\pi]} dt\, \delta\gamma^x(t) \frac{d}{dt} v^x(\gamma(t)) \\ &= -\int_{[0,2\pi]} dt\, \delta\gamma^x(t) \left(\partial_x v^x(\gamma(t))\dot{\gamma}^x(t) + \partial_y v^x(\gamma(t))\dot{\gamma}^y(t)\right).\end{aligned}$$

The x component of (B.2) then takes the form:

$$\begin{aligned}\delta_{\gamma^x} E_{grad}(\gamma) &= -\int_{[0,2\pi]} dt\, \delta\gamma^x(t) \left(\partial_x v^x(\gamma(t))\dot{\gamma}^x(t) + \partial_y v^x(\gamma(t))\dot{\gamma}^y(t)\right) \\ &\quad + \int_{[0,2\pi]} dt\, \delta\gamma^x(t) \left(\partial_x v^x(\gamma(t))\dot{\gamma}^x(t) + \partial_x v^y(\gamma(t))\dot{\gamma}^y(t)\right) \\ &= \int_{[0,2\pi]} dt\, \left[-\delta\gamma^x(t)\partial_y v^x(\gamma(t))\dot{\gamma}^y(t) + \delta\gamma^x(t)\partial_x v^y(\gamma(t))\dot{\gamma}^y(t)\right] \\ &= -\int_{[0,2\pi]} dt\, \delta\gamma^x(t)\dot{\gamma}^y(t) \left[\partial_y v^x(\gamma(t)) - \partial_x v^y(\gamma(t))\right].\end{aligned}$$

Analogously, the y component takes the form:

$$\begin{aligned}\delta_{\gamma^y} E_{grad} &= \int_{[0,2\pi]} dt\, \left[-\delta\gamma^y(t)\partial_x v^y(\gamma(t))\dot{\gamma}^x(t) + \delta\gamma^y(t)\partial_y v^x(\gamma(t))\dot{\gamma}^x(t)\right] \\ &= -\int_{[0,2\pi]} dt\, \delta\gamma^y(t)\dot{\gamma}^x(t) \left[\partial_x v^y(\gamma(t)) - \partial_y v^x(\gamma(t))\right].\end{aligned}$$

The variational derivative of $E_{grad}(\gamma)$ is thus

$$\frac{\delta E_{grad}(\gamma)}{\delta\gamma^i(t)} = -\dot{\gamma}^j(t) \left[\partial_j v^i(\gamma(t)) - \partial_i v^j(\gamma(t))\right],$$

where i and j index the Euclidean coordinates on $\mathbb{R}^2$.

B.2 Model II

Here we detail the calculation of the variational derivative of the image gradient energy term only. The expression for the Gaussian term is computed in the same way.

The image gradient term $E_{grad}(\gamma)$ for the curves parametrized by $(x_0, \delta r(t))$, where $\delta r(t)$ is

a radial variation around a circle of radius r_0 and centred at x_0, is given by

$$\begin{aligned} E_{grad}(\gamma) &= \int_{[0,2\pi]} dt\, n(\gamma(t)) \cdot \nabla I(\gamma(t)) \\ &= \int_{[0,2\pi]} dt\, \dot{\gamma}(t) \cdot v(\gamma(t)) \\ &= \int_{[0,2\pi]} dt[(\dot{\delta r}(t)\cos(t) - (r_0 + \delta r(t))\sin(t))v^x(\gamma(t)) \\ &\quad +(\dot{\delta r}(t)\sin(t) + (r_0 + \delta r(t))\cos(t))v^y(\gamma(t))] \ , \end{aligned} \tag{B.3}$$

taking into account that $\dot{x}_0^x(t) = \dot{y}_0^x(t) = 0$.

Let us develop expressions for the first and the second components of the functional derivative of E given by

$$\frac{\delta E}{\delta \gamma(t)} = \left(\frac{\partial E}{\partial x_0(\gamma)}, \frac{\delta E}{\delta \delta r(t)} \right) \ , \tag{B.4}$$

Defining i and j indexing the Euclidean coordinates in $\mathbb{R}^2$, the derivatives of a curve are given as follows

$$\frac{\partial \gamma^i}{\partial x_0^j(\gamma)} = \delta^i_j \ ;$$

and the derivatives of v:

$$\begin{aligned} \frac{\partial v^x(\gamma(t))}{\partial x_0^j(\gamma)} &= \partial_x v^x \cdot \frac{\partial \gamma^x}{\partial x_0^j}(t) + \partial_y v^x \cdot \frac{\partial \gamma^y}{\partial x_0^j}(t) : \\ \frac{\partial v^x(\gamma(t))}{\partial x_0^x(\gamma)} &= \partial_x v^x(\gamma(t)) = -\partial_x \partial_y I(\gamma(t), \\ \frac{\partial v^x(\gamma(t))}{\partial x_0^y(\gamma)} &= \partial_y v^x(\gamma(t)) = \partial_y^2 I(\gamma(t)) \ ; \end{aligned}$$

and

$$\begin{aligned} \frac{\partial v^y(\gamma(t))}{\partial x_0^j(\gamma)} &= \partial_x v^y \cdot \frac{\partial \gamma^x}{\partial x_0^j}(t) + \partial_y v^y \cdot \frac{\partial \gamma^y}{\partial x_0^j}(t) : \\ \frac{\partial v^y(\gamma(t))}{\partial x_0^x(\gamma)} &= \partial_x v^y(\gamma(t)) = \partial_x^2 I(\gamma(t)), \\ \frac{\partial v^y(\gamma(t))}{\partial x_0^y(\gamma)} &= \partial_y v^y(\gamma(t)) = \partial_y \partial_x I(\gamma(t)) \ . \end{aligned}$$

The components of the derivative of E_{grad} can be written:

$$\begin{aligned}\frac{\partial E_{grad}}{\partial x_0^x(\gamma)} &= \int_{[0,2\pi]} dt(\dot\gamma^x \partial_x v^x + \dot\gamma^y \partial_x v^y)(t) \\ &= \int_{[0,2\pi]} dt(-\dot\gamma^x \partial_x \partial_y I + \dot\gamma^y \partial_x \partial_x I)(t), \qquad \text{(B.5)}\end{aligned}$$

(B.6)

$$\begin{aligned}\frac{\partial E_{grad}}{\partial x_0^y(\gamma)} &= \int_{[0,2\pi]} dt(\dot\gamma^x \partial_y v^x + \dot\gamma^y \partial_y v^y)(t) \\ &= \int_{[0,2\pi]} dt(-\dot\gamma^x \partial_y \partial_y I + \dot\gamma^y \partial_y \partial_x I)(t) \ . \qquad \text{(B.7)}\end{aligned}$$

Thus, we obtain:

$$\frac{\partial E_{grad}}{\partial x_0^i(\gamma)} = \int_{[0,2\pi]} dt\ (-n(\gamma(t))) \cdot \partial_i \nabla I(\gamma(t))) \ .$$

To compute, now, a variational derivative w.r.t. $\delta r(t)$ of the image gradient energy term, we write it as a sum of three terms using the notation $p(t) = \delta r(t)$ to simplify the appearance of the formulas:

$$\begin{aligned}E_{grad}(\gamma) &= \int_{[0,2\pi]} dt\Big[\\ &\quad \dot p(t)\Big(\cos(t) v^x(\gamma(t)) + \sin(t) v^y(\gamma(t))\Big) \qquad \text{(B.8)} \\ &\quad -r_0\Big(\sin(t) v^x(\gamma(t)) + \cos(t) v^y(\gamma(t))\Big) \qquad \text{(B.9)} \\ &\quad -p(t)\Big(\sin(t) v^x(\gamma(t)) - \cos(t) v^y(\gamma(t))\Big)\Big] \ . \qquad \text{(B.10)}\end{aligned}$$

In order to obtain an expression for the derivative of $E_{grad}(\gamma)$ we vary separately every term as detailed here below.

Firstly, we compute the derivative of the x component of $v(\gamma)$.

$$\begin{aligned}\delta_{p(t)} v^x(\gamma(t)) &= \delta_{p(t)} v^x(x_0, p(t)) = \int_{[0,2\pi]} dt \frac{\delta v^x(\gamma(t))}{\delta p(t)} \delta p(t) \\ &= \int_{[0,2\pi]} dt \left[\frac{\partial v^x(\gamma(t))}{\partial \gamma^x(t)} \cdot \frac{\partial \gamma^x(t)}{\partial p(t)} + \frac{\partial v^x(\gamma(t))}{\partial \gamma^y(t)} \cdot \frac{\partial \gamma^y(t)}{\partial p(t)}\right] \delta p(t) \ , \qquad \text{(B.11)}\end{aligned}$$

$$\begin{aligned}\frac{\delta v^x(\gamma(t))}{\delta p(t)} &= \frac{\partial v^x(\gamma(t))}{\partial \gamma^x(t)} \cdot \frac{\partial \gamma^x(t)}{\partial p(t)} + \frac{\partial v^x(\gamma(t))}{\partial \gamma^y(t)} \cdot \frac{\partial \gamma^y(t)}{\partial p(t)} \\ &= \partial_x v^x(\gamma(t)) \cos(t) + \partial_y v^x(\gamma(t)) \sin(t) \ . \qquad \text{(B.12)}\end{aligned}$$

The derivative of the y component of v is calculated in the same way as for the x component and takes the form:

$$\frac{\delta v^y(\gamma(t))}{\delta p(t)} = \partial_x v^y(\gamma(t))\cos(t) + \partial_y v^y(\gamma(t))\sin(t) \quad .$$

Then, the variation of (B.8) is given by

$$\begin{aligned}\delta_{p(t)}\left(\int_{[0,2\pi]} dt\ \dot{p}(t)\cos(t)v^x(p(t))\right) &= \int_{[0,2\pi]} dt\ \cos(t)\delta\dot{p}(t)v^x(p(t)) \\ &\quad + \int_{[0,2\pi]} dt\ \cos(t)\dot{p}(t)\delta v^x(p(t)) \quad .\end{aligned} \tag{B.13}$$

Integrating by parts the first term we obtain:

$$\begin{aligned}&\int_{[0,2\pi]} dt\ \cos(t)\delta\dot{p}(t)v^x(p(t)) \\ &= \cos(t)\delta v^x(p(t))\delta p(t)|_0^{2\pi} - \int_{[0,2\pi]} dt\,\delta p(t)\frac{d}{dt}\left[\cos(t)v^x(p(t))\right] \\ &= -\int_{[0,2\pi]} dt\,\delta p(t)\left[-\sin(t)v^x(p(t)) + \cos(t)\frac{d}{dt}v^x(p(t))\right] \\ &= \int_{[0,2\pi]} dt\,\delta p(t)\sin(t)v^x(p(t)) - \\ &\quad - \int_{[0,2\pi]} dt\,\delta p(t)\cos(t)\left[\partial_x v^x\dot{\gamma}^x(t) + \partial_y v^x\dot{\gamma}^y(t)\right] \quad .\end{aligned} \tag{B.14}$$

Now, developing the above formula taking into account the expressions for the derivatives of the field v, and the derivative of γ developed; and taking the sum with the second term of the variation of (B.14) we find:

$$\begin{aligned}&\delta_{p(t)}\left(\int_{[0,2\pi]} dt\ \dot{p}(t)\cos(t)v^x(p(t))\right) \\ &= \int_{[0,2\pi]} dt\,\delta p(t)\sin(t)v^x(p(t)) \\ &\quad - \int_{[0,2\pi]} dt\,\delta p(t)\cos(t)[\partial_x v^x(\dot{p}(t)\cos(t) - (r_0 + p(t))\sin(t))] \\ &\quad - \int_{[0,2\pi]} dt\,\delta p(t)\cos(t)[\partial_x v^x(p(t))(\dot{p}(t)\sin(t) + (r_0 + p(t))\cos(t))] \\ &\quad + \int_{[0,2\pi]} dt\ \dot{p}(t)\cos(t)\delta p(t)\left[\partial_x v^x(p(t))\cos(t) + \partial_y v^x(p(t))\sin(t)\right] \\ &= \int_{[0,2\pi]} dt\,\delta p(t)\sin(t)v^x(p(t)) - \\ &\quad - \int_{[0,2\pi]} dt\,\delta p(t)\cos(t)(r_0 + p(t))\left[-\partial_x v^x(p(t))\sin(t) + \partial_y v^x(p(t))\cos(t)\right] \quad .\end{aligned} \tag{B.15}$$

The derivative then takes the following form:

$$\frac{\delta\left(\int_{[0,2\pi]} dt\ \dot{p}(t)\cos(t)v^x(p(t))\right)}{\delta p(t)}$$
$$= \sin(t)v^x(p(t)) - \cos(t)(r_0 + p(t))\left[-\partial_x v^x(p(t))\sin(t) + \partial_y v^x(p(t))\cos(t)\right]. \quad \text{(B.16)}$$

Analogously to (B.16),

$$\frac{\delta\left(\int_{[0,2\pi]} dt\ \dot{p}(t)\sin(t)v^y(p(t))\right)}{\delta p(t)}$$
$$= -\cos(t)v^y(p(t)) - \sin(t)(r_0 + p(t))\left[-\partial_x v^y(p)\sin(t) + \partial_y v^y(p)\cos(t)\right]. \quad \text{(B.17)}$$

Then, the variation of (B.10) is given by

$$\delta_{p(t)}\left(\int_{[0,2\pi]} dt\ p(t)(\sin(t)v^x(t))\right) = \int_{[0,2\pi]} dt\ \delta p(t)\sin(t)v^x(p(t)) + \int_{[0,2\pi]} dt\ p(t)\sin(t)\delta v^x(p(t))\ . \quad \text{(B.18)}$$

Thus,

$$\frac{\delta\left(\int_{[0,2\pi]} dt\ p(t)\sin(t)v^x(t)\right)}{\delta p(t)}$$
$$= \sin(t)v^x(p(t)) + p(t)sin(t)\left[-\partial_x v^x(p(t))\cos(t) + \partial_y v^x(p(t))\sin(t)\right]\ ; \quad \text{(B.19)}$$

and analogously

$$\delta_{p(t)}\left(\int_{[0,2\pi]} dt\ cos(t)v^y(p(t))\right) = \int_{[0,2\pi]} dt\ \delta p(t)\cos(t)v^y(p(t)) + \int_{[0,2\pi]} dt\ p(t)cos(t)\delta v^y(p(t))\ , \quad \text{(B.20)}$$

then,

$$\frac{\delta\left(\int_{[0,2\pi]} dt\ \cos(t)v^y(p(t))\right)}{\delta p(t)}$$
$$= \cos(t)v^y(p(t)) + p(t)\cos(t)\left[\partial_x v^y(p(t))\cos(t) + \partial_y v^y(p(t))\sin(t)\right]\ . \quad \text{(B.21)}$$

Now, taking the sum of the above four terms with the appropriate signs, we obtain:

$$(B.16) + (B.17) + (B.19) + (B.21)$$
$$= -\cos(t)\left(-r_0\partial_x v^x(p(t))\sin(t) + (r_0 + p(t)\partial_y v^x(p(t))\cos(t)\right)$$
$$-sin(t)\left(r_0\partial_y v^y(p(t))\cos(t) - (r_0 + p(t)\partial_x v^y(p(t))\sin(t)\right)$$
$$-p(t)\left[-\cos(t)\partial_x v^y(p(t))\cos(t) + sin(t)\partial_y v^x(p(t))\sin(t)\right]\ .$$

The variational derivative of (B.9) is given by

$$\begin{aligned}
&\frac{\delta\left(\int_{[0,2\pi]} dt\; r_0(\sin(t)v^x(p(t)) + \cos(t)v^y(p(t)))\right)}{\delta p(t)} \\
&= \quad r_0 \sin(t)\left[\partial_x v^x(p(t))\cos(t) + \partial_y v^x(p(t))\sin(t)\right] + \\
&\quad\; + r_0 \cos(t)\left[\partial_x v^y(p(t))\cos(t) + \partial_y v^y(p(t))\sin(t)\right] \quad .
\end{aligned}$$

Finally, taking the sum of the two expressions above, the second component of the variational derivative of the image gradient term can be written as follows:

$$\frac{\delta E_{grad}}{\delta p(t)} \quad = \quad (r_0 + p(t))(\partial_x v^y(p(t)) - \partial_y v^x(p(t))) \tag{B.22}$$

Appendix C

Publications and scientific activities

International conferences

- M. S. Kulikova, M. Mani, A. Srivastava, X. Descombes, and J. Zerubia. 'Tree Species Classification Using Radiometry, Texture and Shape Based Features.' In *Proc. of European Signal Processing Conference (EUSIPCO)*, Poznań, Poland, September, 2007.

- M. S. Kulikova, I. H. Jermyn, X. Descombes, E. Zhizhina, and J. Zerubia. 'A Marked Point Process Model with Strong Prior Shape Information for Extraction of Multiple, Arbitrarily-Shaped Objects.' In *IEEE Signal-Image Technology & Internet-Based Systems Conference (SITIS)*, Marakech, Morocco, December, 2009.

- M. S. Kulikova, I. H. Jermyn, X. Descombes, E. Zhizhina, and J. Zerubia. Extraction of Arbitrarily Shaped Objects Using Stochastic Multiple Birth-and-Death Dynamics and Active Contours. In *Proc. of IS&T/SPIE Electronic Imaging Conference*, San Jose, USA, January 2010.

International workshops

- M. S. Kulikova, X. Descombes, and J. Zerubia. 'Shape Recognition for Tree Species Classification.' *EURANDOM workshop on image analysis and inverse problems*, Eindhoven, The Netherlands, 2006.

Seminars

- Seminar on 'Marked point processes for multiple arbitrarily-shaped object extraction' at Florida State University (FSU), USA, February 27, 2009.

- Presentation of the work of thesis at the "Cross Seminar" Ariana/Asclepios/Odyssee at INRIA Sophia-Antipolis, April 7, 2009.

Teaching

- Teaching assistant for 'Games and Strategies' for the first year students at Ecole Polytechnique Nice-Sophia Antipolis, 64h, 2007.
- Teaching assistant for "Document Creation with (X)HTML/XML/LaTeX/MSOffice" for the first year students at Ecole Polytechnique Nice-Sophia Antipolis, 64h, 2008.
- Teaching assistant for "Document Creation with (X)HTML/XML/XSLT/LaTeX/OOffice" for the first year students at Ecole Polytechnique Nice-Sophia Antipolis, 64h, 2009.

Other scientific activities

- Presentation of the work of thesis (poster) at the "Shape Day" workshop, Statistics Department, FSU, USA, April 6, 2007.
- Presentation of the work of thesis (poster) at the summer school ISSPR, Plymouth, UK, July 22, 2007.
- Participation in the French-American INRIA/FSU Associated Team 'SHAPES', [`http://www-sop.inria.fr/ariana/Projets/Shapes/`].
- Participation in the French/Russian/Belorussian INRIA/IITP/UIIP Associated Team 'ODESSA' [`http://www-sop.inria.fr/ariana/Projets/Odessa/`] .

Bibliography

[Bayburt *et al.*, 2008] S. Bayburt, G. Büyüksalih, I. Baz, K. Jacobsen, and Kersten T. Detection of Changes in Istanbul Area With Medium and High Resolution Space Images. In *Proc. of Commission VII, International Society for Photogrammetry and Remote Sensing Congress*, Beijing, 2008.

[Boser *et al.*, 1992] B.E. Boser, I.N. Guyon, and V.N. Vapnik. A Training Algorithm for Optimal Margin Classifiers. In *in Fifth Annual Workshop on Computational Learning Theory*, Pittsburgh, Pennsylvania, United States, 1992.

[Bresenham, 1965] J.E. Bresenham. Algorithm for computer control of a digital plotter. *IBM Systems Journal*, 4:25–30, 1965.

[Burgers, 1998] C.J.C. Burgers. *Tutorial on Support Vector Machines for Pattern Recognition*. Kluwer Academic Publishers, Boston, United States, 1998.

[Caselles *et al.*, 1997] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic Active Contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.

[Chatelain *et al.*, 2009] F. Chatelain, X. Descombes, and J. Zerubia. Parameter Estimation for Marked Point Processes. Application to Object Extraction from Remote Sensing Images. In *Proc. Energy Minimization Methods in Computer Vision and Pattern Recognition*, Bonn, Germany, August 2009.

[Chen *et al.*, 1998] C.H. Chen, L.F. Pau, and P.S. Wang. *The Handbook of Support Vector Machines for Pattern Recognition and Computer Vision (2nd Edition)*. World Scientific Publishing Co., 1998.

[Cremers *et al.*, 2006a] D. Cremers, S.J. Osher, and S. Soatto. Kernel Density Estimation and Intrinsic Alignment for Shape Priors in Level Set Segmentation. *International Journal of Computer Vision*, 69(3):335–351, 2006.

[Cremers *et al.*, 2006b] D. Cremers, N. Sochen, and C. Schnörr. A Multiphase Dynamic Labeling Model for Variational Recognition-Driven Image Segmentation. *International Journal of Computer Vision*, 66(1):67–81, 2006.

[Cremers *et al.*, 2006c] D. Cremers, F. Tischaeuser, J. Weickert, and C. N., Schnörr. Diffusion Snakes: Introducing Statistical Shape Knowledge into the Mumford-Shah Functional. *International Journal of Computer Vision*, 33(3):347–359, 2006.

[Daley and Vere-Jones, 1988] D.J. Daley and D. Vere-Jones. *An introduction to the Theory of Point Processes*. Springer, New-York, 1988.

[Descamps *et al.*, 2009] S. Descamps, X. Descombes, A. Béchet, and J. Zerubia. Détection de Flamants Roses par Processus Ponctuels Marqués pour l'Estimation de la Taille des Populations. *Traitement du Signal*, 28(2), July 2009.

[Descombes *et al.*, 2009] X. Descombes, R. Minlos, and E. Zhizhina. Object Extraction Using a Stochastic Birth-and-Death Dynamics in Continuum. *Journal of Mathematical Imaging and Vision*, 33:347–359, 2009.

[Erikson, 2004a] M. Erikson. *Segmentation and Classification of Individual Tree Crowns*. PhD thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden, 2004.

[Erikson, 2004b] M. Erikson. Species Classification of Individual Segmented Tree Crowns in High Resolution Aerial Images using Radiometric and Morphologic Image Measures. *Remote Sensing of Environment*, 91:469–477, 2004.

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

[Gougeon *et al.*, 1998] F.A. Gougeon, D.G. Leckie, I. Scott, and D. Paradine. Individual Tree Crown Species Recognition: the Nahmint Study. In D.A. Hill and D.G. Leckie, editors, *Proc. of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*, pages 209–223, Pacific Forestry Center, Victoria, British Columbia, Canada, February 1998.

[Gougeon, 1995a] F.A. Gougeon. Comparison of Possible Multispectral Classification Schemes for Tree Crown Individually Delineated on High Spatial Resolution MEIS Images. *Canadian Journal of Remote Sensing, 21(1)*, pages 1–9, 1995.

[Gougeon, 1995b] F.A. Gougeon. A Crown-following Approach to the Automatic Delineation of Individual Tree Crowns in High Spatial Resolution Aerial Images. *Canadian Journal of Remote Sensing, 21(3)*, pages 274–284, 1995.

[Gougeon, 1996] F.A. Gougeon. Vers l'Inventaire Forestier Automatisé : Reconnaître l'Arbre ou la Forêt ? CD-ROM du 9^{eme} Congrès de L'Association Québécoise de Télédétection, Mai 1996.

[Green, 1995] P.J. Green. Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. *Biometrika 82*, pages 711–732, 1995.

[Haralick, 1978] R.M. Haralick. Statistical and Structural Approaches to Texture. In *Proc. IEEE*, 1978.

[Horvath, 2007] P. Horvath. *The 'Gas of Circles' Model and Its Application to Tree Crown Extraction*. PhD thesis, Universite de Szeged, Universite de Nice Sophia Antipolis, December 2007.

[Isard and Blake, 1998] M. Isard and A. Blake. Condensation-Conditional Density Propagation for Visual Tracking. *International Journal of Computer Vision*, 29(1):5–28, 1998.

[Jacobsen, 2006] M. Jacobsen. *Point Process Theory and Applications : Marked Point and Piecewise Deterministic Processes*. Birkhäuser, Boston, 2006.

[Joshi and Srivastava, 2009] S. H. Joshi and A. Srivastava. Intrinsic Bayesian Active Contours for Extraction of Object Boundaries in Images. *International Journal of Computer Vision*, 81(3):331–355, 2009.

[Joshi *et al.*, 2007a] S. Joshi, E. Klassen, A. Srivastava, and I. H. Jermyn. A Novel Representation for Riemannian Analysis of Elastic Curves in $\mathbb{R}^n$. In *Proc. Computer Vision and Pattern Recognition*, Minneapolis, USA, June 2007.

[Joshi *et al.*, 2007b] S. Joshi, E. Klassen, A. Srivastava, and I. H. Jermyn. Removing Shape-Preserving Transformations in Square-Root Elastic (SRE) Framework for Shape Analysis of Curves. In *Proc. Energy Minimization Methods in Computer Vision and Pattern Recognition*, Ezhou, China, August 2007.

[Joshi, 2007] S. Joshi. *Inferences in Shape Space with Applications to Image Analysis and Computer Vision*. PhD thesis, The Florida State University, Thallahassee, Florida, USA, 2007.

[Juan *et al.*, 2006] O. Juan, R. Keriven, and G. Postelnicu. Stochastic Motion and the Level Set Method in Computer Vision: Stochastic Active Contours. *International Journal of Computer Vision*, 69(1):7–25, 2006.

[Kass *et al.*, 1998] M. Kass, A. Wiktkin, and D. Terzopolous. Shakes: Active Contours Models. *International Journal of Computer Vision*, 60:173–195, 1998.

[Kervann and Heitz, 1998] C. Kervann and H. Heitz. A Hierarchical Markov Modelling Approach for the Segmentation and Tracking of Deformable Shapes. *Graphical Models and Image Processing*, 60:173–195, 1998.

[Khudeev, 2005] R. Khudeev. A New Flood-Fill Algorithm for Closed Contours. In *Proc. IEEE Conference on Control and Communications*, 2005.

[Klassen *et al.*, 2004] E. Klassen, A. Srivastava, W. Mio, and S.H. Joshi. Analysis of Planar Shapes Using Geodesic Paths on Shape Spaces. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26:372–383, 2004.

[Kulikova *et al.*, 2007] M. S. Kulikova, M. Mani, A. Srivastava, X. Descombes, and J. Zerubia. Tree Species Classification Using Radiometry, Texture and Shape Based Features. In *Proc. European Signal Processing Conference*, 2007.

[Kulikova *et al.*, 2010] M. S. Kulikova, I. H. Jermyn, X. Descombes, E. Zhizhina, and J. Zerubia. Extraction of Arbitrarily Shaped Objects Using Stochastic Multiple Birth-and-Death Dynamics and Active Contours. In *Proc. IS&T/SPIE Electronic Imaging*, San Jose, USA, January 2010.

[Lacoste *et al.*, 2005] C. Lacoste, X. Descombes, and J. Zerubia. Point Processes for Unsupervised Line Network Extraction in Remote Sensing. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 27(10):1568–1579, 2005.

[Lacoste, 2004] C. Lacoste. *Extraction de Réseaux Linéiques à partir d'Images Satellitaires et Aériennes par Processus Ponctuels Marqués*. PhD thesis, Universite de Nice Sophia Antipolis, September 2004.

[Leckie *et al.*, 2003] D.G. Leckie, F.A. Gougeon, N. Walsworth, and D. Paradine. Stand Delineation and Composition Estimation using Semi-Automated Individual Tree Crown Analysis. *Remote Sensing of Environment, 85*, pages 355–369, 2003.

[Leventon and Grimson, 2000] M.E. Leventon and and Faugeras O. Grimson, W.E.L. Statistical Shape Influence in Geodesic Active Contours. In *Proc. Computer Vision and Patter Recognition*, Hilton Head Island, SC, USA, 2000.

[Miller and Pankov, 2001] B.M. Miller and A. R. Pankov. *Theory of Stochastic Processes in Examples and Problems (in Russian)*. Moscow Aviation Institute, Moscow, 2001.

[Ortner *et al.*, 2007] M. Ortner, X. Descombes, and J. Zerubia. Building Outline Extraction from Digital Elevation Models Using Marked Point Processes. *International Journal of Computer Vision*, 72(2):107–132, 2007.

[Ortner, 2004] M. Ortner. *Processus Ponctuels Marqués pour l'Extraction Automatique de Caricatures de Bâtiments à partir de Modèles Numériques d'Elévation*. PhD thesis, Universite de Nice Sophia Antipolis, October 2004.

[Osher and Fedkiw, 2003] S. Osher and R. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer Verlag, 2003.

[Perrin *et al.*, 2005] G. Perrin, X. Descombes, and J. Zerubia. A Marked Point Process Model for Tree Crown Extraction in Plantations. In *Proc. IEEE International Conference on Image Processing*, Genova, 2005.

[Perrin *et al.*, 2006] G. Perrin, X. Descombes, and J. Zerubia. A Non-Bayesian Model for Tree Crown Extraction Using Marked Point Processes. *Research Report 5846, INRIA, France*, 2006.

[Perrin, 2006] G. Perrin. *Etude du Couvert Forestier par Processus Ponctuels Marqués*. PhD thesis, Ecole Doctorale de Centrale Paris, Paris, France, 2006.

[Preston, 1977] C. Preston. Spatial Birth-and-Death Processes. *Bulletin of the International Statistical Institute*, 46(2):371–391, 1977.

[Rathi *et al.*, 2007] Y. Rathi, N. Namrata Vaswani, A. Tannenbaum, and A. Yezzi. Tracking Deforming Objects Using Particle Filtering for Geometric Active Contours. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29(8):1470–1475, 2007.

[Rousson and Paragios, 2002] M. Rousson and N. Paragios. Shape Priors for Level Set Representations. In *Proc. 7th European Conference on Computer Vision-Part II*, pages 78–92, London, UK, 2002. Springer-Verlag.

[Sethian, 1999] J.A. Sethian. *Level Set Methods and Fast Marching Methods.* Cambridge Press, 1999.

[Stoica *et al.*, 2000] R. Stoica, X. Descombes, and J. Zerubia. Road Extraction in Remotely Sensed Images Using a Stochastic Geometry framework. In *Proc. IEEE International Workshop Bayesian Inference and Maximum Entropy Methods*, Gif-sur-Yvette, France, 2000.

[Storvik, 1994] G. Storvik. A Bayesian Approach to Dynamic Contours Through Stochastic Sampling and Simulated Annealing. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(10):976–986, 1994.

[Stroock, 2005] D.W. Stroock. *An Introduction to Markov Processes*. Berlin: Springer, 2005.

[Tu *et al.*, 2002] Z. Tu, S. C. Zhu, and H.-Y. Shum. Image Segmentation by Data Driven Markov Chain Monte Carlo. In *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pages 131–138, 2002.

[van Lieshout, 1993] M.N.M. van Lieshout. Stochastic Annealing for Nearest-Neighbour Point Processes with Application to Object Recognition. *Advances in Applied Probability*, pages 281–300, 1993.

[van Lieshout, 2000] M.N.M. van Lieshout. *Markov Point Processes and Their Applications*. World Scientific Publishing Company, 2000.

[Vapnik, 1995] V.N. Vapnik. *The Nature of Statistical Learning Theory*. Springer Verlag, New-York, 1995.

[Vapnik, 1998] V.N. Vapnik. *Statistical Learning Theory*. Wiley-Interscience, 1998.

[Younes, 1998] L. Younes. Computable Elastic Distance Between Shapes. *SIAM Journal of Applied Mathematics*, 58:565–586, 1998.

Résumé

Les travaux réalisés lors cette thèse s'inscrivent dans le cadre de l'analyse d'images, appliquée aux problèmes de télédétection. Nous nous sommes interessés à deux sujets principaux: 'L'étude et la classification d'espèces d'arbres', décrit dans la première partie du manuscript, et 'L'extraction d'objets multiples de forme arbitraire par Processus Ponctuels Marqués', présenté dans la seconde partie.

Les techniques de télédétection sont aujourd'hui largement développées et trouvent de nombreuses applications en foresterie où l'information obtenue par photointerprétation des photographies aériennes est combinée avec les mesures effectuées sur le terrain. Ces techniques permettent d'étudier la biodiversité de l'écosystème forestier. Les méthodes d'analyse d'images de forêts ont pour but de faciliter la tâche d'inventaire forestier. Les paramètres d'importance relevés sur les images et le terrain sont la densité de peuplement, l'âge des arbres, le volume des troncs, la composition d'espèces d'arbres, et des valeurs écologiques comme le biotope et l'habitat. Afin d'obtenir des informations sur la diversité d'espèces de forêts par exemple, la classification des couronnes d'arbres est nécessaire. Radiométrie et texture sont deux critères largement utilisés dans les méthodes de classification existantes. Nous proposons, dans cette thèse, d'utiliser des informations supplémentaires obtenues à partir de l'étude de la forme des couronnes (*i.e.* houppiers) d'arbres afin de les classifier. Les formes considerées sont des courbes planes fermées représentant les bords des couronnes d'arbres extraites d'images aériennes infrarouge couleur (IRC) de résolution 3cm/pixel, fournies par l'Université des Sciences Agricoles de Suéde. La résolution allant jusqu'au niveau de détail des feuilles des arbres, l'information géométrique fournie rend donc intéressante l'utilisation de la forme des couronnes.

Les images à partir desquelles les couronnes des arbres sont extraites représentent les zones forestières où les arbres sont proches du nadir, *i.e.* les couronnes d'arbres sont vues de dessus presque verticalement. Les données IRC fournissent plus d'information sur la végétation que les couleurs naturelles. Cela permet de distinguer par la teinte les feuillus des résineux ainsi que les arbres sains des autres, car l'image IRC est très dépendante de l'activité chlorophyllienne de la plante. Une fois numérisées, les photos IRC sont représentées par des couleurs artificielles pour décrire les trois bandes [PIR (Proche InfraRouge), Rouge, Vert]: le rouge de l'image représente le PIR; le vert représentant le rouge et le bleu

le vert.

La classification est effectuée sur les quatre espèces d'arbres les plus répandues en Suède: l'épicéa, le pin sylvestre, le bouleau et le tremble. Deux de ces espèces sont des conifères et les deux autres sont des feuillus. Pour nos expérimentations, nous avons sélectionné 48 couronnes (12 par classe). Leur contours, représentés par un ensemble ordonné de points, ont été délimités manuellement afin de préserver les détails de leur forme qui est étudiée par la suite.

Pour nos études, nous avons choisi une méthodologie fondée sur l'analyse de formes de courbes fermées continues dans un espace de formes en utilisant la notion de chemin géodésique. Nous réalisons d'abord des études utilisant la métrique obtenue avec une représentation de la courbe par la fonction d'angle développée par Klassen *et al.* [Klassen *et al.*, 2004], puis en utilisant la métrique élastique et la représentation par la racine-carrée appelée q-fonction proposée par Joshi *et al.* [Joshi *et al.*, 2007a, Joshi *et al.*, 2007b]. Cette méthodologie a été choisie pour plusieurs raisons. La forme des couronnes d'arbres à partir d'images de résolution 3cm/pixel a une structure complexe. L'hypothèse de continuité permet d'éviter le choix de points de référence ('landmarks') sur la courbe. Les espaces de formes avec des métriques appropriées sont alors définis comme des variétés non-linéaires utilisant les propriétés d'invariance de la forme: *e.g.* translation, rotation rigide, dilatation uniforme, changement de point de départ le long de la courbe, et enfin la re-paramétrisation dans le cas de la métrique élastique. A chaque chemin géodésique dans l'espace de formes est associée une distance. Nous avons évalué cette distance en tant que mesure de similarité des formes de couronnes d'arbres dans le sens d'appartenance à l'une ou l'autre espèce.

Les résultats obtenus en utilisant la métrique non-élastique montrent que la similarité de deux arbres au sens de la distance géodésique ne reflète pas l'appartenance à une même espèce. L'étude des formes intermédiaires sur le chemin géodésique montre que la transformation passe souvent par une forme plus étendue (cette forme intermédiaire est proche de celle d'un bouleau). Par conséquent, la distance géodésique utilisant la métrique non-élstique ne peut être utilisée en tant que tel pour la classification des arbres selon les espèces. Cela est dû à la nature de la métrique qui ne permet que de tordre la courbe en la transformant en une autre courbe le long du chemin géodésique, ce qui signifie qu'il est moins coûteux de déplier la courbe, puis de la plier de nouveau. Ceci implique qu'une forme régulière comme celle du bouleau ou du pin devrait être plus proche que celle du tremble ou de l'épicéa.

La seconde métrique utilisée pour l'étude de la forme des couronnes d'arbres, est fondée sur l'analyse de Riemann des courbes élastiques. La métrique élastique initialement proposée par Younes [Younes, 1998], permet à une courbe de s'étirer ainsi que de se plier en se déformant le long d'une géodésique dans l'espace de formes correspondant. Les expériences faites en utilisant la métrique élastique montrent que la distance géodésique entre deux

arbres d'une même espèce est plus petite que celle de deux arbres d'espèces différentes, dans le cas où les couronnes des arbres étudiées ont une forme très caractéristique de l'espèce. Mais la forme de la couronne varie beaucoup au sein d'une même classe. Ceci peut être dû, par exemple, à l'âge de l'arbre: un jeune pin a une forme acérée similaire à celle de l'épicéa, ou un épicéa peut avoir une couronne de branches dense, ce qui entraîne une forme circulaire similaire à des formes de couronnes des pins. Une autre expérience a été faite pour les couronnes d'arbres d'espèces différentes mais de formes similaires. Les résultats obtenus montrent que la similarité de formes est aussi reflétée par la distance géodésique.

Des études effectuées, nous avons donc conclu que la forme seule n'est pas suffisamment discriminante pour la classification des arbres, en raison de la complexité et de la variété de la forme au sein d'une même espèce. Toutefois, elle peut être utile en complément de l'information sur la radiométrie et la texture.

L'étape suivant de la thèse a donc été d'inclure l'information de la forme des couronnes des arbres dans un processus de classification, afin de quantifier l'amelioration de la performance d'un classifieur utilisant les descripteurs de radiométrie et de texture. La classification a été effectuée avec une méthode de classification supervisée en utilisant un séparateur à vaste marge, avec un noyau gaussien. Premièrement, la classification a été réalisée en utilisant uniquement les caractéristiques radiométriques. Ensuite, les caractéristiques de texture ont été ajoutés et enfin, des descripteurs de forme ont été inclus. La performance a été calculée après chaque étape. Les performances de chacune des étapes ont été évaluées via les taux moyen P et maximum P_{max} des couronnes bien classées.

Dans notre étude, l'IRC permet de distinguer entre conifères et feuillus. Les quatre classes d'arbres - le tremble, le bouleau, l'épicéa et le pin - peuvent donc être facilement identifiées comme feuillus ou conifères à partir des statistiques du premier ordre (la moyenne et l'écart type calculés à partir de l'histogramme des intensités des pixels de l'image), $P = 0.54$ et $P_{max} = 0.67$. Pour distinguer encore davantage au sein des classes de feuillus et de conifères, une analyse de texture a été réalisée en utilisant une matrice de co-occurrence (GLCM) de niveau de gris. Une matrice de co-occurrence est une représentation quantitative en deux dimensions de la relation spatiale [Haralick, 1978], bien adaptée pour la caractérisation des micro-textures. L'information de texture a permis d'améliorer les résultats de la classification pour les feuillus, $P = 0.71$ et $P_{max} = 0.833$. Cela est dû au fait que l'une des espèces de feuillus (bouleau) a une texture à peu près uniforme contrairement aux trembles.

Ensuite, une analyse de forme de la couronne des arbres a été effectuée afin de déterminer les informations qui pourraient nous permettre de classer les espèces. Les couronnes d'un tremble, par exemple, ont une structure irrégulière, où les convexités correspondants aux branches sortant du corps des couronnes sont importantes. Les branches des épicéas sont plus régulières et sont orientées radialement. Les contours de la couronne des bouleaux

et des pins sont plus circulaires, celles des bouleaux étant les plus lisses. Néanmoins, il faut noter, que malgré les différences décrites, un nombre important de contours restent inclassables par analyse visuelle en raison de la variabilité au sein d'une même classe. La représentation par la fonction d'angle semble bien adaptée pour exploiter les différences de forme des couronnes. A partir de la fonction d'angle, un ensemble de descripteurs de forme est inclu dans le vecteur des descripteurs. A l'aide de nos observations sur la représentation par la fonction d'angle, nous notons, en particulier, que les géodésiques passent toujours par des formes circulaires et que la forme des couronnes de pin et de bouleau est plus proche du cercle, que celle du tremble ou du sapin. La première caractéristique choisie comme descripteur est donc une distance géodésique à un cercle calculée dans l'espace de forme avec la représentation de forme par la fonction d'angle. Puis, nous traduisons la propriété que les couronnes de certaines espèces ont une structure plus régulière que les autres, par une mesure de l'élasticité des courbes. Par ailleurs, les convexités (branches, feuilles) des couronnes d'épicéas sont généralement moins nombreuses et plus grandes que les convexités de couronnes de bouleau. Ce critère se reflète dans le nombre de maxima locaux de la fonction d'angle. La forme de la couronne des pins est assez proche de celle du bouleau, mais certaines irrégularités sont plus grandes, avec quelques branches qui dépassent. Nous quantifions donc les irrégularités des contours de la couronne en fonction de la taille de l'irrégularité (représetant des branches, des feuilles et des ombres). En intégrant des caractéristiques de forme, la performance de la classification se traduit par une amélioration d'environ 4%, $P = 0.747$, alors que le rendement maximal atteint 87.5%. Ainsi, la forme a permis d'améliorer la performance de la classification en espèces au sein des classes de conifères et de feuillus.

Afin de classifier les couronnes des arbres à partir d'images, leur détection et leur segmentation à partir de ces images est nécessaire. A la résolution des images que nous avons utilisées pour nos études, leur géométrie est clairement visible (les arbres feuillus étant identifiables presqu'au niveau de la feuille), et devrait donc être prise en compte pour associer un arbre à une espèce. En outre, les zones forestières représentent des scènes très complexes difficiles à segmenter. Ceci est dû au fait que les arbres sont en général situés à approximité les uns des autres, certains arbres ayant des couronnes qui se chevauchent. Le défi principal est donc le suivant : les objets qui se chevauchent doivent être segmentés comme un ensemble d'objets individuels et non comme un objet seul. C'est une tâche particulièrement difficile pour les arbres regroupés qui appartiennent à la même espèce (même pour l'extraction manuelle par un expert). Nous nous sommes donc intéressés à développer une méthode qui permet d'extraire automatiquement de multiples objets à partir d'images à très haute résolution, les images représentant des scènes complexes contenant des objets qui se superposent dans l'image.

Les processus ponctuels marqués offrent une technique d'extraction d'objets multiples reconnue pour sa capacité à inclure de l'information sur la géométrique des objets recherchés, ainsi que sur leurs interactions. Plusieurs modèles de processus ponctuels marqués ont été

appliqués avec succès au problème de l'extraction d'objets à partir d'images de haute résolution, les objets ayant une géométrie visible simple [Lacoste *et al.*, 2005, Perrin *et al.*, 2005, Ortner *et al.*, 2007, Descamps *et al.*, 2009]. Le deuxième sujet de cette thèse, présenté dans la seconde partie, consiste donc en une généralisation de l'approche par processus ponctuels marqués pour l'extraction d'objets multiples à forme complexe arbitraire à partir d'images à très haute résolution.

Pour ce faire, une solution aurait été de définir la distribution sur l'espace de toutes les courbes fermées, mais les espaces de tels éléments sont des variétés non-linéaires à la topologie complexe. La définition d'une métrique appropriée sur ces espaces n'est pas simple. Par conséquent, même la description d'une dynamique simple comme la diffusion pour un objet, devient un problème difficile. Pour pallier cette difficulté, nous avons proposé de réduire l'espace d'objets à un sous-ensemble d'objets, adaptés localement aux données. Chaque objet est donc un minimum local de l'énergie associée à un contour actif. Nous utilisons, ensuite, l'énergie totale comprenant des interactions multiples pour trouver la configuration optimale composée de ces objets, localement adaptés à l'image. L'espace d'objet considéré reste donc de petite dimension, mais les objets possibles individuels sont déterminés non pas a priori, mais par les données de l'image ainsi que par un modèle lié à l'objet individuel. En conséquence, les objets sont des courbes fermées dans le domaine de l'image. Une fois l'espace objet défini, nous considèrons une énergie de Gibbs, et donc une distribution de probabilité sur l'espace des configurations composées d'un nombre arbitraire d'objets. Cette énergie se décompose en la somme des énergies des objets individuels de la configuration, plus un terme d'interaction qui contrôle la superposition des objets. L'énergie d'un seul objet est la somme d'un terme de lissage de la courbe, et d'un terme d'attache aux données reliant un objet à l'image. Pour trouver la configuration optimale d'objets multiples, on calcule une estimation du Maximum A Posteriori (MAP) par échantillonnage à partir d'un processus de naissances et morts multiples incorporé dans un schéma de recuit simulé.

En l'absence d'un terme d'interaction, l'estimation du MAP serait alors composée du sous-ensemble des objets ayant une énergie négative. Cela conduirait sans doute à trouver une configuration qui serait une solution dégénérée. Voila pourquoi un terme d'interaction est ajouté. Il contrôle la relation entre les objets différents, en particulier il décourage le recouvrement.

De manière plus détaillée, le modèle proposé se décrit comme suit: les différents objets sont définis par l'évolution d'un certain nombre de courbes initiales par une technique de descente de gradient vers les minima locaux de l'énergie liée à l'objet individuel. Les courbes initiales sont des cercles centrés en un pixel de l'image, avec des rayons compris dans un certain intervalle, le rayon étant la marque associée à chaque point. La taille de l'espace d'objet individuel n'est donc pas supérieure à celle de l'ensemble des courbes initiales. Ceci est similaire aux ensembles d'objets utilisés dans les modèles de processus ponctuels

précédents préservant ainsi l'efficacité des calculs. Pour trouver la configuration optimale d'objets multiples, durant l'étape de naissance nous plaçons, via un processus uniforme de Poisson, un certain nombre de centres de cercles initiaux avec des rayons uniformément échantillonnés. Les cercles sont ajoutés à la configuration courante (vide au départ), et sont ensuite, déformés pour atteindre un minimum local de l'énergie de chaque objet, produisant ainsi une configuration d'objets multiples adaptés localement à l'image. L'étape de mort supprime, ensuite, un certain nombre d'objets avec une probabilité qui dépend de la température et la différence des énergies (pondérées par la température) entre les configurations avec et sans chacun des objets dans la configuration. Le processus itératif est stoppé si tous les objets ajoutés à l'étape de naissance, et seulement ceux-ci, sont retirés à l'étape suivante de mort. La performance de l'approche est démontrée par les résultats expérimentaux sur des données synthétiques et réelles.

Une série d'expériences de différents niveaux de complexité a été faite pour tester le modèle. A partir des résultats obtenus, nous avons conclu que ce modèle est bien adapté aux scènes composées d'objets qui ne varient pas trop en forme et en taille dans une classe, et qui ont des bords assez lisses. Dans ce cas, la précision du nombre d'objets extraits peut être excellente si l'on favorise le terme de lissage des bords, contrôlant ainsi la longueur de la courbe représentant les bords de l'objet. Le modèle peut donc, dans une certaine mesure, séparer des objets distincts qui se chevauchent, mais cela conduit à une imprécision dans la délimitation des objets. La solution que nous proposons pour résoudre ce problème est d'incorporer une information de forme a priori plus spécifique dans l'énergie de l'objet individuel afin de détecter des objets dans des scènes d'images de haute complexité contenant des objets se superposant.

Le second objectif de la seconde partie de cette thèse est donc une nouvelle extension du cadre des processus ponctuels marqués, qui consiste à définir une méthode pour intégrer une information a priori forte sur la forme des objets recherchés, sans augmenter la dimensionnalité de l'espace objet (et donc la complexité du calcul) dans le modèle défini précédemment. Les principes restent les mêmes, mais l'espace d'objet individuel est défini de manière à tenir compte des connaissances a priori sur la forme des objets que l'on souhaite extraire.

L'ensemble des objets possibles est donc défini à partir des minima locaux d'une énergie qui intègre des informations provenant des données et l'information a priori de la forme des objets à extraire. Cela permet de traiter des configurations composées d'objets aux formes complexes et qui se superposent. L'utilisation de l'information de forme a priori présente un autre avantage: même si l'objet à extraire est en partie obscurci ou occulté, par exemple, par une ombre, comme les couronnes des épicéas, l'objet peut néanmoins être correctement extrait comme un seul objet jusqu'à un certain degré de précision.

Pour résumer les travaux présentés dans la seconde partie de la thèse, nous avons effectué

une comparaison des modèles utilisant ou non l'information a priori des différentes formes sur une image à très haute résolution. Cette comparaison montre que le modèle de processus ponctuels marqués d'objets simples permet l'extraction rapide des objets à partir d'une image, mais la précision géométrique reste faible pour les objets aux formes complexes. Par conséquent, ce type de modèle est un outil approprié pour la détection rapide d'objets dans les images à plus basse résolution. La seconde approche, un modèle de processus ponctuels marqués pour l'extraction d'objets à forme arbitraire avec seulement l'a priori de forme faible (terme de lissage de la courbe), est géométriquement bien plus précise, tout en n'augmentant pas la complexité du calcul indûment. Les limites de cette approche sont cependant dues au fait qu'elle inclut seulement une information faible sur la forme pour définir les objets individuels possibles, elle ne permet pas de détecter différents types d'objets ayant les mêmes caractéristiques radiométriques, ni de séparer deux ou plusieurs objets qui se superposent dans l'image. Le modèle de processus ponctuels marqués pour l'extraction d'objets à forme arbitraire avec un a priori de forme fort permet de s'affranchir de ces deux limitations, comme le montre les résultats obtenus sur les images de zones forestières naturelles de très haute résolution.

ABSTRACT

This thesis includes two main parts. In the first part we address the problem of tree crown classification into species using shape features, without, or in combination with, those of radiometry and texture, to demonstrate that shape information improves classification performance. For this purpose, we first study the shapes of tree crowns extracted from very high resolution colour aerial infra-red images. For our study, we choose a methodology based on the shape analysis of closed continuous curves on shape spaces using geodesic paths under the bending metric with the angle-function curve representation, and the elastic metric with the square root q-function representation. A necessary preliminary step to classification is extraction of the tree crowns. In the second part, we address thus the problem of extraction of multiple objects with complex, arbitrary shape from remote sensing images of very high resolution. We develop a model based on marked point processes. Its originality lies in its use of arbitrarily-shaped objects as opposed to parametric shape objects, *e.g.* ellipses or rectangles. The shapes considered are obtained by local minimisation of an active contour energy with weak and then strong shape prior knowledge included. The objects in the final (optimal) configuration are then selected from amongst these candidates by a multiple birth-and-death dynamics embedded in an annealing scheme. The approach is validated on very high resolution images of forest provided by the Swedish University of Agriculture.

Keywords: colour infra-red aerial image, classification, tree crown, shape, multiple object extraction, marked point process, birth-and-death dynamics, active contour, shape prior.

RÉSUMÉ

Cette thèse est composée de deux parties principales. La première partie est dédiée au problème de la classification d'espèces d'arbres en utilisant des descripteurs de forme, en combainison ou non, avec ceux de radiométrie ou de texture. Nous montrons notamment que l'information sur la forme améliore la performance d'un classifieur. Pour ce faire, dans un premier temps, une étude des formes de couronnes d'arbres extraites à partir d'images aériennes, en infrarouge couleur, est effectuée en utilisant une méthodologie d'analyse de formes des courbes continues fermées dans un espace de formes, en utilisant la notion de chemin géodésique sous deux métriques dans des espaces appropriés : une métrique non-élastique en utilisant la reprèsentation par la fonction d'angle de la courbe, ainsi qu'une métrique élastique induite par une représentation par la racine-carée appelée q-fonction. Une étape préliminaire nécessaire à la classification est l'extraction des couronnes d'arbre. Dans une seconde partie, nous abordons donc le problème de l'extraction d'objets de forme complexe arbitraire, à partir d'images de télédétection à très haute résolution. Nous construisons un modèle fondé sur les processus ponctuels marqués. Son originalité tient dans sa prise en compte d'objets de forme arbitraire par rapport aux objets de forme paramétrique, *e.g.* ellipses ou rectangles. Les formes sélectionnées sont obtenues par la minimisation locale d'une énergie de type contours actifs avec différents a priori sur la forme incorporé. Les objets de la configuration finale (optimale) sont ensuite sélectionnés parmi les candidats par une dynamique de naissances et morts multiples, couplée à un schéma de recuit simulé. L'approche est validée sur des images de zones forestières à très haute résolution fournies par l'Université d'Agriculture de Suède.

Mots clefs: image aérienne infrarouge couleur, classification, couronne d'arbre, forme, extraction d'objets multiples, processus ponctuel marqué, dynamique de naissances et morts, contour actif, a priori de la forme.