

Variational Methods for Texture Segmentation

THÈSE N° 4303 (2009)

PRÉSENTÉE LE 27 FÉVRIER 2009

À LA FACULTÉ SCIENCES ET TECHNIQUES DE L'INGÉNIEUR

LABORATOIRE DE TRAITEMENT DES SIGNAUX 5

PROGRAMME DOCTORAL EN INFORMATIQUE, COMMUNICATIONS ET INFORMATION

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Nawal HOUHOU

acceptée sur proposition du jury:

Prof. A. Skrivervik Favre, présidente du jury

Prof. J.-Ph. Thiran, directeur de thèse

Dr X. Bresson, rapporteur

Dr S. Makram-Ebeid, rapporteur

Prof. P. Vanderghyest, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2009

Contents

Contents	i
List of Figures	v
Remerciements	vii
Abstract	ix
Version Abregée	xi
Notations	xiii
1 Introduction	1
1.1 The Image Segmentation Problem	1
1.2 Segmentation Methods	3
1.3 Energy based Segmentation Methods	3
1.4 Motivations and Contributions of this Thesis	5
1.5 Organization of this Thesis	6
I Feature Definition: An Essential Step Before Segmenting	7
2 Feature for Region-Based Segmentation	11
2.1 Texture Features	11
2.2 Region Measures for Feature Analysis	13
2.2.1 First and Second Order Statistical Moment	15
2.2.2 Probability Estimation	15
2.3 Conclusion	16
3 A New Intrinsic Geometric Texture Descriptor based on the Principal Curvatures	17
3.1 Textures and Beltrami Framework	17

3.2	New Texture Descriptor based on the Beltrami Framework and the Shape Operator	19
3.3	Results	20
3.4	Conclusion	21
II	Unsupervised Segmentation	27
4	Background: Active Contours and Information Theory	31
4.1	Active contours	31
4.1.1	Boundary-Based Active Contours	31
4.1.2	Region-Based Segmentation	35
4.2	Global Minimization of Active Contour Energy	36
4.2.1	Model 1: Global Active Contours from Partial Differential Equation	36
4.2.2	Model 2: Global Active Contours from the Variational Model	38
4.2.3	Model 1 vs Model 2	38
4.2.4	Numerical Scheme	39
4.3	Information Theoretic and Probabilistic tools	39
4.3.1	Shannon Entropy	40
4.3.2	Relative Entropy	40
4.3.3	From Image Space to Probability Space	40
4.4	Conclusion	41
5	Unsupervised Region Competition Based on Active Contour and Kullback-Leibler Distance	43
5.1	Proposed 2-Phase Segmentation Method	43
5.2	Existence of Minimizers for the Proposed Model	44
5.2.1	Our Image Segmentation Model	45
5.2.2	Generalization of the Existence Theorem	47
5.2.3	47
5.3	Dual Formulation and Fast Segmentation	48
5.4	Experimental Results	49
5.4.1	Gray-Value Feature	49
5.4.2	Intrinsic Geometric Feature	49
5.5	Discussion	53
6	Global Multi-Phase Segmentation	55
6.1	Multi-Region Segmentation with Level Set Function	55
6.2	Our New Global Multi-Phase Method	57
6.2.1	Four Regions Case	57
6.2.2	Generalization for a Global Solution and Fast Algorithm based on Dual Formulation	58
6.2.3	Existence of a Global Solution	59
6.3	Results	60
6.4	Conclusion	61

III Non-local Semi-Supervised Segmentation Based on Continuous Mincut	65
7 Introduction to Mincut and Semi-supervised Segmentation	69
7.1 Graph Diffusion	69
7.1.1 Image Modeled by a Weighted Graph	69
7.1.2 Mincut or the Diffusion on the Graph	71
7.2 Semi-Supervised Segmentation	73
7.3 Conclusion	74
8 Non-local Semi-Supervised Segmentation based on Continuous Mincut	75
8.1 Proposed Segmentation Method	75
8.1.1 Continuous Min Cut (CMC)	75
8.1.2 Semi-Supervised Segmentation with Labels	76
8.1.3 Our Semi-Supervised Segmentation Algorithm based on Mincut . . .	77
8.2 Study of the Model	78
8.3 Conclusion	79
9 Results and Discussion	81
9.1 Experimental Results	81
9.1.1 Test on the Parameters	81
9.1.2 TV-Regularization Effect	84
9.1.3 Texture Images	84
9.1.4 Natural Images	84
9.1.5 Color Images	85
9.1.6 Medical Images	85
9.2 Discussion and Conclusion	85
10 Conclusion	91
10.1 Achievements	91
10.2 Future Work	92
A Annex	93
A.1 Shape Derivation Tool	93
Bibliography	95
Curriculum Vitae	103

List of Figures

1.1	Example of natural and medical images.	2
1.2	Example of textures from natural and medical images.	2
2.1	Gabor response for a pulsation.	14
3.1	Beltrami Framework	18
3.2	Features extraction of a natural image.	23
3.3	Synthetic non-oriented textured image and its pdf	24
3.4	Structure tensor of Figure 3.3(b).	24
3.5	Gabor filter responses of figure 3.3(b)	25
3.6	Our texture descriptor on Figure 3.3(b).	25
3.7	Failure example for our texture descriptor.	26
4.1	The level-set function	34
5.1	Segmentation results based on KL distance for synthetic noisy image	50
5.2	Unsupervised Segmentation results based on KL distance for natural images	51
5.3	Unsupervised Segmentation results based on KL distance and our texture descriptor for textured images	51
5.4	Unsupervised Segmentation results based on KL distance and our texture descriptor for natural images	52
5.5	Unsupervised Segmentation results based on KL distance and our texture descriptor for natural images	53
6.1	The Vese-Chan approach for four-phase case	57
6.2	Four-phase Segmentation of synthetic and real-world image	60
6.3	Four-phase Segmentation of synthetic and real-world textural images with region competition based on pdf	62
6.4	[Four-phase Segmentation of a 3D volume of the brain	63
7.1	Bridges of Konisberg. Figure from [34]	70
9.1	(a)Synthetic image test. (b)Initialization with label L_1	82
9.2	Semi-Supervised Segmentation: effect of parameter cl	83
9.3	Semi-Supervised Segmentation: effect of parameter a	83
9.4	Semi-Supervised Segmentation: effect of parameter f	83

9.5	Semi-Supervised Segmentation: effect of parameter σ_2	84
9.6	Effect of TV-norm on the Semi-Supervised Segmentation	85
9.7	Semi-Supervised Segmentation for synthetic textured images	86
9.8	Semi-Supervised Segmentation for natural images	88
9.9	Semi-Supervised Segmentation for color images	89
9.10	Semi-Supervised Segmentation for 2D medical images	90

Remerciements

Au terme de ces quatre années passées au Laboratoire des Traitements des Signaux pour accomplir ce travail de thèse, je tiens à remercier ceux qui m'ont permis d'y arriver.

Mes premiers remerciements vont à mon directeur de thèse, Prof. Jean-Philippe Thiran pour m'avoir accueilli dans son groupe et permis de réaliser ce travail dans les meilleures conditions qui soient. Merci pour ta confiance, ton optimisme et ton soutien.

Je tiens également à remercier le Dr. Arnaud Garcia qui a été le responsable de mon projet de master et qui m'a mise en contact avec Jean-Philippe.

Merci aux membres de mon jury de thèse, à la présidente du jury, Prof. Anja Skrivervik Favre d'avoir mené les débats lors de ma soutenance privée. Merci au Dr. Sherif Makram-Ebeid pour toutes ses observations et ses critiques constructives. Merci au Prof. Pierre Vandergheynst pour les discussions intéressantes.

Le travail présenté ici est en réalité le résultat d'une fructueuse collaboration avec le Dr. Xavier Bresson. Merci Xavier de m'avoir tant enseigné avec pédagogie, gentillesse et patience. Merci également à Karen, Olympe et toi pour votre chaleureux accueil et les bons moments passés ensemble lors de mon séjour en Californie.

Si la vie d'un étudiant aux LTS est si paisible c'est bien parce que Madame Marianne Marion s'occupait de tous nos soucis administratifs. Merci Marianne pour ton professionnalisme, ton efficacité mais surtout pour ta bonne humeur. Merci à Rosy Depietro d'avoir pris la relève avec efficacité. Merci également à Gilles Auric pour l'assistance logistique et informatique.

Un grand merci à ma post-doc, le Dr. Meritxell Bach Cuadra pour avoir relu une grande partie de ma thèse mais également pour tous les conseils et aides tout le long de ces quatre années, en particulier pour les présentations orales. Merci également au Dr. Leila Cammoun et à Subrahmanyam Gorthi pour les collaborations dans le domaine médical.

Une grande partie de ce projet s'est faite sous forme de programmation informatique. Merci au Dr. Valérie Duay pour m'avoir initié aux mystères d'ITK et surtout à Ruth Campos qui a toujours su résoudre tous mes problèmes (informatiques) tout au long de ces quatre années.

L'Institut des Traitements des Signaux (ITS) est un endroit propice à l'enrichissement scientifique mais aussi culturel. Merci au Prof. Murat Kunt de nous avoir offert ce cadre dynamique et multiculturel.

Merci à tous les membres des ITS qui se sont succédés tout au long de ces quatre années pour l'atmosphère chaleureuse et amicale. En particulier, au tupper club, Meri, Elisa, Val, Uigi et Ruth pour leur amitié, au club baby foot, Julien, Matteo, Philippe, Mathieu, Yann,

Ulrich, David, Anna...et tous les autres, et aux habitués des pauses café Leila, Lisa, Patricia et Florian. Je tiens à remercier mes collègues de bureau, Florent pendant une première période suivi de Ruth. C'était réellement un plaisir de partager le bureau avec chacun d'entre vous.

Enfin, je remercie Meri et Dani, ma famille adoptive à Lausanne, de votre attention et votre amitié.

Evidemment tout ceci n'aurait été possible sans le soutien de ma famille, mes parents, mes frères et soeur. Merci à mes parents d'avoir fait depuis toujours de nos études et notre réussite leur priorité.

Abstract

In the last decades, image production has grown significantly. From digital photographs to the medical scans, including satellite images and video films, more and more data need to be processed. Consequently the number of applications based on digital images has increased, either for medicine, research for country planning or for entertainment business such as animation or video games. All these areas, although very different one to another, need the same image processing techniques.

Among all these techniques, *segmentation* is probably one of the most studied because of its important role. Segmentation is the process of extracting meaningful objects from an image. This task, although easily achieved by the human visual system, is actually complex and still a true challenge for the image processing community despite several decades of research.

The thesis work presented in this manuscript proposes solutions to the image segmentation problem in a well established mathematical framework, i.e. variational models. The image is defined in a continuous space and the segmentation problem is expressed through a functional or energy optimization. Depending on the object to be segmented, this energy definition can be difficult; in particular for objects with ambiguous borders or objects with textures. For the latter, the difficulty lies already in the definition of the term *texture*. The human eye can easily recognize a texture, but it is quite difficult to find words to define it, even more in mathematical terms. There is a deliberate vagueness in the definition of texture which explains the difficulty to conceptualize a model able to describe it. Often these textures can neither be described by homogeneous regions nor by sharp contours.

This is why we are first interested in the extraction of texture features, that is to say, finding one representation that can discriminate a textured region from another. The first contribution of this thesis is the construction of a texture descriptor from the representation of the image similar to a surface in a volume. This descriptor belongs to the framework of non-supervised segmentation, since it will not require any user interaction. The second contribution is a solution for the segmentation problem based on active contour models and information theory tools. third contribution is a semi-supervised segmentation model, i.e. where constraints provided by the user will be integrated in the segmentation framework. This processus is actually derived from the graph of image patches. This graph gives the connectivity measure between the different points of the image. The segmentation will be expressed by a graph partition and a variational model.

This manuscript proposes to tackle the segmentation problem for textured images.

Keywords:Variational Methods, Segmentation, Active Contours, Texture, Mincut.

Version Abregée

La production d'images a été l'apanage de ces dernières décennies. De la photographie numérique au scan medical, en passant par les images satellitaires et les films vidéo, la quantité de données à traiter est toujours de plus en plus importante. Le nombre de domaines qui en résultent a naturellement augmenté, que ce soit pour des applications médicales, de recherche pour l'aménagement du territoire, ou simplement des applications dans le domaine du divertissement, tels que les animations ou les jeux vidéos. Tous ces domaines, bien que très différents, font souvent appel aux mêmes techniques de traitement de l'image.

Parmi ces techniques, la *segmentation* est sans doute l'une des plus étudiées, de par l'importance de son rôle. La segmentation est le processus d'extraction des différents objets constituant une image. Cette tâche bien que relevée avec facilité par le système visuel humain, est en réalité complexe et reste un véritable défi pour la communauté du traitement de l'image malgré plusieurs décennies de recherche.

Le travail de thèse présenté dans ce manuscrit propose des solutions aux problèmes de segmentation d'images dans un cadre mathématique solide, à savoir les modèles variationnels. L'image est alors définie dans un espace continu et le problème de segmentation est exprimé à travers l'optimisation d'une fonctionnelle appelée énergie. Selon l'objet à segmenter, la définition de cette énergie peut s'avérer difficile; en particulier pour les objets qui présentent des frontières ambiguës avec l'objet voisin ou des objets présentant des textures. Pour ces dernières, la difficulté commence déjà dans la définition même du terme *texture*. Une texture est facile à reconnaître par l'œil humain mais assez difficile à définir avec des mots, d'autant plus en termes mathématiques. En effet, il existe un flou artistique autour de ce terme qui entraîne une difficulté à conceptualiser un modèle qui puisse les décrire. Souvent ces textures ne peuvent être décrites ni par des régions homogènes ni par des contours nets. C'est la raison pour laquelle nous nous sommes d'abord intéressé à l'extraction de caractéristiques pour les textures, c'est-à-dire trouver une représentation qui puisse discriminer une région texturée d'une autre. La première contribution de cette thèse est un descripteur de texture, construit à partir d'une représentation de l'image comme une surface dans un volume. Ce descripteur sera alors intégré dans le cadre d'une segmentation non-supervisée, c'est-à-dire une segmentation libre de toute interaction avec un utilisateur. La deuxième contribution proposée dans cette thèse, est donc une solution pour la segmentation basée sur les modèles des contours actifs et sur des outils de la théorie de l'information.

La troisième contribution est un modèle de segmentation semi-supervisée, c'est-à-dire que des contraintes fournies par l'utilisateur seront intégrées dans le processus de segmenta-

tion. Ce processus est en réalité dérivé d'une représentation par un graphe de parcelles de l'image. Le graphe donne une mesure de connectivité entre les différents points de l'image. La segmentation sera traduite à travers la partition du graphe par un modèle variationnel.

Ce manuscrit propose donc de s'attaquer au problème de la segmentation d'images présentant des zones texturées.

Mots Clefs: Méthodes Variationnelles, Segmentation, Texture, Contours Actifs, Min-cut.

Notations

Abbreviations

2D	two dimensional
3D	three dimensional
GAC	Geodesic Active Contour
KL	Kullback-Leibler
LS	Level Set
pdf(s)	probability density function(s)
TV	Total Variation
Mincut	Minimal Cut

Miscellaneous

Ω	Open subset of \mathbb{R}^N
$\partial\Omega$	Boundary of Ω
$ \Omega $	Area of Ω

Functions and Functionals

$\frac{\partial u}{\partial \zeta}, \partial_\zeta u, u_\zeta$	Differentiation of the function u w.r.t. the variable ζ
$\chi_\Omega, \mathbf{1}_\Omega$	Characteristic function of the set $\Omega \subset \mathbb{R}^n$ s.t $\chi_\Omega = 1$ if $x \in \Omega$ and 0 otherwise
$\frac{\partial F}{\partial u}$	First derivation of the functional F w.r.t. the function u called the Euler-Lagrange equation
\mathbf{x}	point in \mathbb{R}^n s.t. $x = (x : 1, \dots, x_N)$
$d\mathbf{x}$	Lebesgue measure in \mathbb{R}^n
$TV(u)$	Total variation of the function u s.t. $TV(u) = \int_{\mathbb{R}^n} \nabla u $

1

Introduction

1.1 The Image Segmentation Problem

Dealing with information extracted from a natural image, a medical scan, satellite data or a frame in a video sequence is the purpose of image analysis.

In the real world, the stimulus that is received by the retina is perceived as a whole and complete information. Between the electromagnetic reception and the perception, physiological and neurological processes construct the final perception and analysis of the image. In fact vision is composed of many interacting components including analysis of color, texture and shape, the whole conducted by prior knowledge of the human brain. Computer vision aims at getting the same result as human perception. The computer interface receives the image as a matrix of pixels/voxels and several levels of processes are involved to get, when it is possible, the same result as human analysis. The collection of processes involved in the visual perception are usually hierarchically classified as belonging to either *low level vision* or *high level vision*. High level vision consists on the interpretation of the image following some rule or prior knowledge. In low level vision, image processing is performed to extract some visible physical properties in the image such as shape and boundaries or to improve the quality of the image. In this thesis we will be dealing with image processing and more precisely with the *image segmentation* task. The objective of *segmentation* methods is to determine a *partition of an image into a finite number of semantically important regions* such as anatomical or functional structures in medical images or objects in natural images. The segmentation task has been studied for several decades, however it is still a challenging task. This task is essential in many applications including face detection in video sequences, changes detection in satellite images, anatomical or functional object extraction in medical images or object extraction in natural images.

Example on the two last applications will be shown all along this thesis. Figure 1.1 shows the type of images which will be used to illustrate the efficiency of new algorithms proposed in this thesis. In those examples we typically want to extract the zebra from

its background (Figure 1.1(a)), the liver from Computed Tomography (CT) scan of the abdomen (Figure 1.1(b)) or the white and gray matter matter in the brain Magnetic Resonance Image (MRI) (Figure 1.1(c)).

We are in particular interested on images that present textured regions. At this stage of the manuscript, we should set straight the signification of the word *texture*. In fact we can all recognize a texture when we see one but a proper mathematical definition is still difficult to give. However, it is consensually admitted that textures are fine scale-details, usually with some periodicity and oscillatory nature [4]. All along this work, the term *texture* groups all the images that are composed partially or totally of textured part. Example of textures are given on Figure 1.2. Texture images can be encountered in different kind of domain and the partition of the different objects in these images is an important issue in many applications.

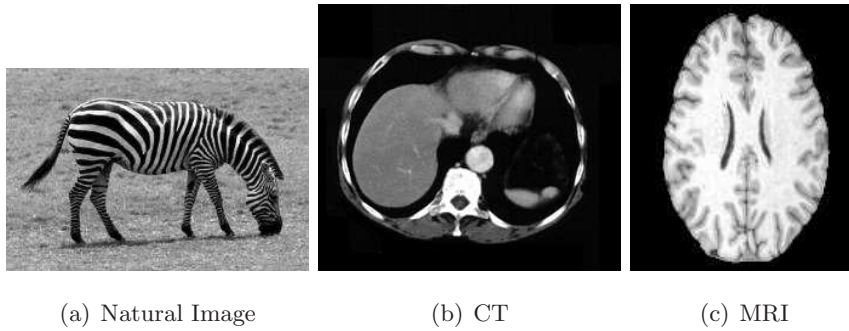


Figure 1.1: Example of natural and medical images that will be handled in this thesis for the segmentation task.

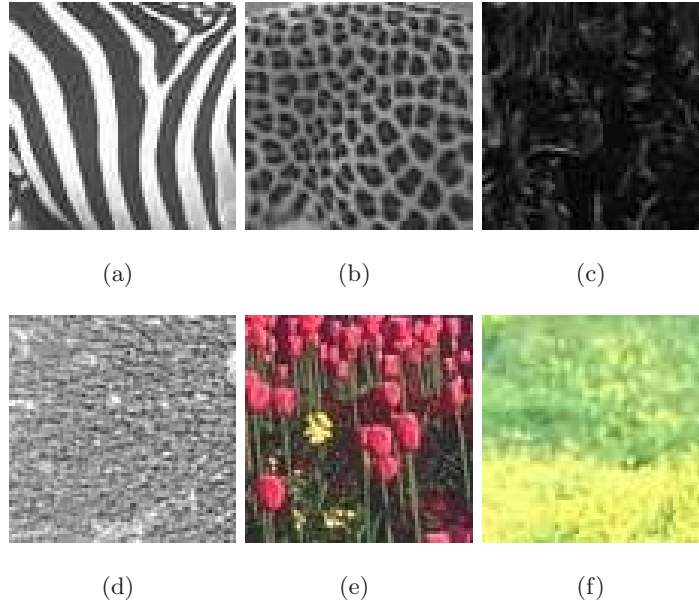


Figure 1.2: Example of textures from natural and medical images. (a) and (b) Patterns from animals. (c) Curly hairs. (d) Liver section (e) Flowers (f) Grass.

1.2 Segmentation Methods

Many approaches have been proposed to solve the image segmentation problem. A review of the huge number of segmentation methods is beyond the scope of this thesis. The book of Sonka et al. [95] provides a nice overview of some of these segmentation methods. The most intuitive and simple one is probably image thresholding. In fact, several objects can be described by homogeneous intensity and it can be possible to extract the object from it background by fixing a constant threshold value. Usually a study of the image histogram helps to find the right threshold. Segmentation can also be performed by edge detection [58, 83]. Defining sharp and closed contours is a way to detect objects and thus solving the segmentation problem. Edges are detected by filters based on gradient of the image and various techniques allows to have a continuous definition of the object contours [19]. However obtaining sharp and closed contours from a single edge detector is a difficult task, especially for noisy or texture images. In fact the gradient operator is a local operator and this is quite restrictive for images that need to be analyzed in a higher scale. Region based methods allows to develop techniques that are more based on the local homogeneity of a region as region growing techniques [14, 29]. These techniques are based on binary criterion of belonging to a region. This criterion will be used either to region merging, region splitting or a combination of the two. Edge-based and region-based segmentation can also be unified in the same framework. This goal can be achieved with *energy based methods*. These methods can combine both edge-based and region-based method by an elegant mathematical formulation. These energy based methods will be our main interest in this manuscript.

1.3 Energy based Segmentation Methods

Energy based methods formulate the image segmentation problem into an optimization problem such that the equilibrium state corresponds to the segmentation solution. The problem can be expressed in various forms depending on the data nature. Solving image segmentation problem with energy based methods implies to solve the following optimization problem:

$$\min_{\zeta} F(\zeta), \quad (1.1)$$

where F is some continuous and differentiable energy functional, ζ can be a closed contour or/and a piece-wise constant approximation of the image or/and a region. Translating a segmentation problem into an energy functional minimization one allows to incorporate different constraints, i.e. different assumptions on the model are gathered together with a weighted influence. In general the energy imposes to ζ some fidelity to the original image and some regularization behavior.

Energy-based image segmentation resolution is carried on by two principal communities. The first one belongs to the *variational methods* family. In this community, the first assumption is that the image is a continuous function in \mathbb{R}^n , where n is the dimension of the image. Calculus of variation (see Chapter 4 Section 4.1.1) or shape derivative tools (see Annex A.1) are used to find the associated Euler-Lagrange evolution equation which lead to the segmentation solution. Our work belongs to this community.

The second community treats the image in the discrete space i.e the image is digital and

belongs to \mathbb{Z}^n . These methods are combinatorial methods and relies often on graph theory (Chapter 7).

Variational Methods Two main categories of variational methods for image segmentation exist: The Mumford-Shah functional and the active contour (AC) method. Mumford and Shah proposed in 1989 [71] the minimization of a functional of (u, K) where u is a piece-wise constant approximation of the original image I and K is the set of discontinuities, i.e the border between the different regions. Solving the Mumford-Shah Functional is equivalent to find a piece wise constant partition of the image while regularizing the borders between the regions. This optimization problem is not trivial mainly because u and K are not defined in the same space. Several authors proposed regularization solution of this problem [41, 66, 68].

The second category of variational model for image segmentation are Active Contours (AC) or snakes and have been introduced by Kass et al. in [55]. In this image segmentation method, an initial contour is subjected to a speed term which drives toward the boundary of the object to be segmented. The speed term is determined by the minimization of a functional composed of a contour smoothing term (internal energy) and an attraction term of the contour towards the object boundaries (external energy). Based on the method proposed by Kass et al., Caselles et al. and Kichenassamy et al. in [20, 56] proposed the geometrically intrinsic model of geodesic/geometric active contours where the evolution curve is handled by the level set method introduced by Osher and Sethian in [73]. Finally, we notice that authors in [20] reformulated the snake problem as a minimal path in a Riemannian space. The first generation of active contour segmentation methods has been based on edge detection [20, 55, 56]. However, boundary-based segmentation are often too restrictive in many applications such as in medical image segmentation where fuzzy contours can be encountered or natural images with texture where edges usually do not represent the region of interest. To deal with these problems, segmentation models based on region descriptors such as mean, variance, probability density function (pdf) have been developed.

Chan and Vese have proposed in [22] a 2-phase segmentation method based on the mean descriptor where the active contour evolves in such a way that the difference between the inside (resp. outside) gray level value and the inside (resp. outside) mean value is minimized. This model is a particular case of the Mumford and Shah model when the image approximation is two regions piece-wise constant function.

Other statistical moments such as the variance descriptor can be used to carry out the segmentation task (see e.g [53, 107]) but the probability density function is a more general descriptor.

Zhu and Yuille in [110] and Paragios, Rousson and Deriche in [74, 84] approximated the pdf of the given image by a mixture of Gaussians, each one representing a homogeneous intensity region to be segmented. More recently and related to our work, Jehan-Besson, Aubert, Barlaud, Faugeras and Herbulot in [3, 48, 53] used the pdf of evolving regions of interest in given images as a general region descriptor. The regions of interest are given by minimization of region-based functions using the "shape derivative tool" defined by Delfour and Zolesio in [30] (see Annex A.1). Aubert, Barlaud, Faugeras and Jehan-Besson proved in [3] the equivalence of minimizing a region-based functional with the shape derivative tool and minimizing boundary-based functional with the calculus of variations. They applied in [3, 53] the shape derivation tool to the image segmentation problem. Along with the same tradition, Herbulot et al. in [48] used the shape derivative tool and information theory

concepts (entropy and mutual information) to perform image segmentation as well. All these methods allows to find local minimizers of the segmentation problem. In 2004, Bresson et al. [13] proposed a method where the active contour leads to a global solution.

Combinatorial Methods Combinatorial methods solve the optimization problem of a functional F represented in the discrete space \mathbb{Z}^n , where n is the dimension of the image. Let $G = (V, E)$ be a graph where the set V represents the graph nodes (points in the image) and E the edges connection nodes also called arcs connection. Two particular nodes are defined in this graph: node s represents the source of the graph and t the sink. In computer vision, a graph is defined in order to express a cost function or an energy which performs a given task as restoration, segmentation or stereo reconstruction. Minimizing the cut in the graph, i.e finding the best partition of the graph at the least cost, is the goal of the process. The resolution of the minimum cut in a graph can then be solved by using algorithms for maximizing the flow, as for instance the max flow algorithm of Ford and Fulkerson [37]. We will note also the fact that a constraint of the existence of s and t clears the way naturally for semi-supervised segmentation. Greig et al. [46] were the first to use the minimal cut/maximum flow algorithm for image restoration. Wu and Leahy [106] proposed to use the minimal cut algorithm for data clustering. They opened the way to several graph partitioning method. Then Boykov and Jolly [8] adapted the work in [46] for image segmentation. A particularly efficient and fast optimization of the min cut algorithm was proposed by Boykov and Kolmogorov in [9]. In summary the mincut algorithm solves several computer vision problems via the minimization of energy represented by a graph, including image restoration [51, 106], image segmentation [7, 8], stereo and motion [11], texture synthesis [59]. Finally we will notice that graph cut has been linked to active contour with the work in [10]. After defining the appropriate graph which describe the same energy, the active contour optimization is solved by the min cut algorithm.

1.4 Motivations and Contributions of this Thesis

The goal of this thesis is to study the problem of object segmentation and to propose algorithms to solve it. Among all the families of methods that achieve the image segmentation task, energy based methods solved with variational models are the one that has been chosen to introduce new algorithms. The principal reason is that such approaches offer a rigorous mathematical framework. Moreover being in the continuous space, the pixel/voxel resolution is not a limit. The segmentation method that we will propose should be general enough to handle any kind of images including images with textures.

When we tackle the texture image segmentation problem, the problem of extracting pertinent feature arise naturally. Our first motivation is to define a texture feature descriptor. This feature descriptor will be integrated in the segmentation framework. This brings us to the second motivation which is a segmentation framework for two-phase images completely unsupervised, that means independent from the user. Thus, an unsupervised segmentation method that can handle multiple phases. Unsupervised methods offer the comfort of a totally user independent and automatic process will be proposed. However some applications require a user interaction. This is the reason why a semi-supervised segmentation method is also designed. We are motivated by the use of non-local information represented by graphs

and user defined constraints on the object to segment.

Therefore, the main contributions presented in the thesis are fourfold:

- An intrinsic texture descriptor based on a Beltrami representation of the image which allows the use of differential geometry tools like principal curvatures(Chapter 3).
- A global active contour unsupervised region-based segmentation method based on the Kullback-Leibler distance(Chapter 5)
- A global active contour unsupervised region-based method for multi-phase segmentation(Chapter 6)
- A non-local semi-supervised segmentation based on continuous mincut method(Chapter 8)

1.5 Organization of this Thesis

This thesis is composed of three parts. Each part is preceded by a preview of the motivations and the contributions. Then, the reader will find a first introductory chapter where the related state of the art is developed. Each part is self-contained and can be read independently even if some results can be used from one part to another. The first part develops a new descriptor for texture images. This descriptor will be used in the second part. In this second part, a new unsupervised active contour segmentation method is proposed. Finally, the third part presents our original semi-supervised segmentation method based on continuous min cut on the graph. More precisely:

- **Part I: Chapter 2** presents the texture feature representation from the state of the art segmentation literature, followed by region measures to treat these descriptors. **Chapter 3** develops our new texture descriptor followed by a discussion on the position of our texture descriptor with respect to some state of the art descriptors.
- **Part II: Chapter 4** presents the state of the art for active contour methods and introduces tools of information theory and probability. **Chapter 5** contains the development of our proposed unsupervised region based segmentation method. In Section 5.1 the core of our method is presented followed by a mathematical justification in Section 5.2. Then in Section 5.3 a dual formulation which leads to a convex functional and a fast numerical scheme is presented. In Section 5.4 the experimental results on various data type are presented. Finally, **Chapter 6** presents an unsupervised multi-region segmentation method. The segmentation is based on a convex energy functional and its minimization provides global solution.
- **Part III: Chapter 7** is an introduction to the graph-based method for image processing and particularly for the mincut method. An overview of the semi-supervised segmentation in the literature is also given. In **Chapter 8**, the construction of our semi-supervised method is presented. We start in Section 8.1 by proposing a continuous form of the graph partitioning minimal cut. In Section 8.1.2, labels on some part of the objects are added. Apart from the fact that the user have an extra control on the segmentation, it allows to have a proper segmentation model. Section 8.2 studies the mathematical properties of the model. Then in **Chapter 9**, the experimental results on different kind of data are shown, followed by a discussion on the model.

Part I

Feature Definition: An Essential Step Before Segmenting

Motivation and Contribution

Motivation The general idea of this part is to define a feature descriptor in the goal to fulfil the region-based image segmentation task. The first assumption in region-based segmentation is that each region regroups the same feature. It seems thus essential to have a pertinent feature on which the segmentation task can be based. Naturally the quality of the segmentation is highly dependent on the quality of the features. In fact following the nature and complexity of the data, a preprocessing to extract the feature which can allow the segmentation has been proved to be necessary. For instance, the raw textured images are often too weak to be incorporated directly in the segmentation process. Synthetic and natural images containing textured regions are our main interest. This part aims at defining an interesting feature for textured images and a descriptor measures which can be efficient for the segmentation task.

Contribution The contribution of this part is a texture descriptor based on a Beltrami representation of the image which allows the use of differential tools.

Feature for Region-Based Segmentation

2

In this chapter some state-of-the-art on feature descriptors is first described. Then, in Section 2.2, global measures that can discriminate different regions are defined.

2.1 Texture Features

In image processing, a feature on a raw image is some relevant characteristic of the image; relevant in the sense that it can be used in an automatic process for a given task. Features can represent edges, corners or homogeneous regions. We are interested in the extraction of homogeneous region features.

Feature extraction for the segmentation purpose is a subject of interest for several decades [72]. This research area is wide and covers field as motion descriptor (for ex. optical flow), boundary descriptors (for ex. edge detection operators) or texture descriptor. Here, we are interested in this last field of research.

There is a deliberate vagueness in the definition of texture probably because texture is a concept peculiar to human perception. This implies that there is neither a unique definition nor unique mathematical model of this concept. Thus defining a "universal" texture descriptor is a difficult task.

The first and most natural texture feature is the image itself. In an image $I \subset \mathbb{R}^2$ each pixel is characterized by its gray-value or intensity. If the image is composed of multiple channels (as color images) then each pixel has a vector of intensity which characterizes it. When the image is composed of textured region, the pixel intensities values do not give any pertinent information. In fact textures can not be analyzed at the pixel scale but need to be analyzed at higher level scale, scale where information on the neighborhood around the pixel is taken into account. In fact, besides of the intensity information, texture descriptor should consider the scale and the orientation (or lack of orientation) of the image. A natural approach for texture segmentation is to first represent the texture image by feature descriptors and then to apply a vector-valued segmentation scheme. It is clear that the quality of

the segmentation will depend on the extraction of good features for region discrimination.

Among texture descriptors, there are the ones that are statistical based as the co-occurrence matrix. A co-occurrence matrix describes how often gray-values between two pixels, with a certain distance at a given direction, occurs. This way the gray-level configuration is built. The configuration varies rapidly with distance in fine textures and slowly for coarse textures [47]. Several tools offer a statistical study of the texture as for example edge frequency, primitive length approaches (see [95] and references therein) or fractal methods [78].

An other class of method for texture feature extraction are the transform-based approaches of the image. Filters as for instance the gradient filter or the wavelet bank filter [63] have been used for texture feature extraction, in particular before the image segmentation task [85, 87]. The structure tensor which is based on the gradient of the image and the Gabor filters which belong to the wavelet filters family will be developed hereafter. In fact these two descriptors are used in segmentation methods related to the one that we will propose in the second part of this thesis.

Structure Tensor The structure tensor has been first introduced in [38] for corner detection, then it has been used to deal with oriented textures in [81]. The classical structure tensor is the result of the tensorial product between the smoothed gradient version of a function u and its transpose:

$$\begin{aligned} u_\sigma &= K_\sigma * u \\ ST(\nabla u_\sigma) &= \nabla u_\sigma \nabla u_\sigma^T, \end{aligned}$$

where K_σ is the Gaussian Kernel with standard deviation σ .

For an image $I \in L^\infty(\Omega)$

$$ST(\nabla I_\sigma) = K_\sigma * \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}, \quad (2.1)$$

where I_x and I_y are the derivative of the image I in respectively the horizontal and vertical direction. The structure tensor is an attractive tool because it allows orientation description and image structure analysis [105]. However, the Gaussian smoothed gradient of an image suffers of one principal problem: dislocation of the edges and the structures. To cope with this problem, it is now popular to use non-linear version of the structure tensor by applying an isotropic or anisotropic diffusion. In [17], Brox et al. demonstrate the efficiency of the non-linear structure tensor in comparison to the classical one. Then in [16], Brox and Weickert proposed a modified version of the structure tensor where locale scale estimation based on the region size is integrated.

Gabor Filter Bank Texture features generated by Gabor/Morlet wavelet transform [60] are powerful tools to discriminate textures of different orientation and scale. The main motivation is based on the fact that simple cells in the visual cortex can be modeled by Gabor functions [64]. The Gabor functions are parameterized by a wavelet orientation angle θ and a scale σ (Figure 2.1). Following [97], the Gabor function can be modeled by a complex sinusoidal signal centered at the 2D frequency coordinates (ω_x, ω_y) and modulated by a Gaussian envelope :

$$\Psi(x, y) = g'(x, y) \exp(j(\omega_x x + \omega_y y)),$$

where

$$g'(x, y) = \frac{1}{\lambda\sigma^2} g\left(\frac{x'}{\lambda\sigma}, \frac{y'}{\sigma}\right), g(x, y) = \frac{1}{2\pi} g\left(-\frac{x^2 + y^2}{2}\right),$$

and

$$\begin{aligned} x' &= x \cos \theta + y \sin \theta, \\ y' &= -x \sin \theta + y \cos \theta. \end{aligned}$$

λ is the aspect ratio between x and y scales, σ is the scale parameter, and θ the orientation parameter of the wavelet. $g'(x, y)$ is a Gaussian function spatially scaled σ and rotated by θ

For a dirac pulsation, we can see in Figure 2.1 the magnitude Gabor response for four different orientations ($\theta = \{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\}$) and scales. Given a certain number of orientations and scales, the original image can be reconstructed from Gabor filter responses obtained by convolution of the given image and the set of (θ, σ) -parameterized Gabor functions. Obviously, increasing the number of orientations and scales will improve the reconstruction quality. However, a good reconstruction can also be achieved by selecting only the most relevant filter responses [52].

The Gabor filter will be used to compare and evaluate the feature descriptor that will be proposed in the Chapter 3.

In the next section region measures are defined. These measures will be used on the feature of interest for the segmentation task.

2.2 Region Measures for Feature Analysis

Once the texture is represent by the appropriate feature vector, measurements characterizing the difference between regions need to be defined.

Since textures show certain consistent properties, one way to describe such textures or its features is through their statistical properties. In order to extract statistical region descriptor, the intensity value of the image $I(x)$ or the vector intensity of the feature vector $F(I(x))$ is considered as a random variable with a distribution f_I . We are then able to compute statistical information on a region $\Omega \subset \mathbb{R}^n$.

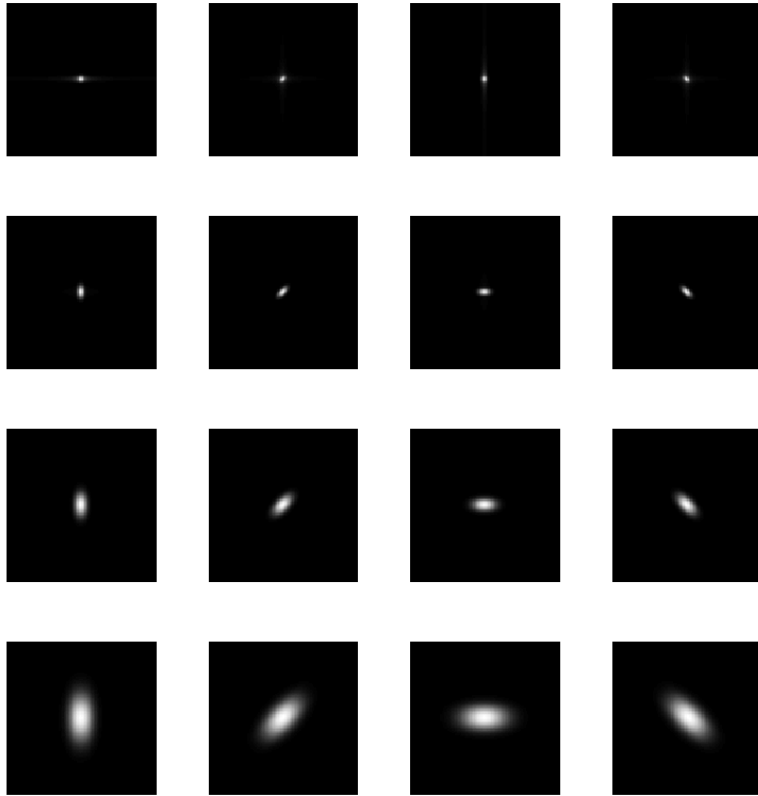


Figure 2.1: Gabor response for a pulsation for 4 different orientation ($\theta = \{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\}$) and scale

2.2.1 First and Second Order Statistical Moment

The average μ of the region is the first natural descriptor:

$$\mu = \sum_{\Omega} f_I I(x).$$

If f_I is uniformly distributed, i.e the probability of having any $I \in [0, p-1]$, p is the number of bins, is equiprobable then we can express the average or mean as follows:

$$\mu = \frac{1}{|\Omega|} \sum_{\Omega} I(x).$$

This descriptor is very efficient for piece-wise constant or piece-wise smooth image [22] if computed on the gray-level image, or for texture images computed on a vector of Gabor response [87].

The variance of the region can also be used as a descriptor and is defined as:

$$\sigma^2 = \frac{1}{|\Omega|} \sum_{\Omega} (I(x) - \mu)^2.$$

This descriptor has been used for image segmentation purpose in [53, 107].

2.2.2 Probability Estimation

The distribution of a sample is probably one of the most efficient way to describe a region. It provides a complete description of the variation in a sample which is in our case the intensity image or features distribution. This distribution can be estimated in a parametric or non-parametric way via a probability density function (pdf).

Parametric Estimation Parametric models make strong assumptions on the pdf of the regions. In image processing, a gaussian distribution is often chosen. Modeling the intensity I distribution over a region Ω as a Gaussian distribution can be expressed as follows:

$$p_G(I) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(I-\mu)^2/2\sigma^2},$$

where μ and σ are respectively the mean and variance parameters which must be estimated from the samples.

Zhu and Yuille in [110] and Paragios, Rousson and Deriche in [74, 84] approximated the pdf of the image by a mixture of Gaussians, each one representing a homogeneous intensity region to be segmented. For medical image analysis this assumption is often used and had proved its efficiency as for instance in the work of Bach et al. [28]. A review of parametric based segmentation methods can be found in [27].

Non-Parametric Estimation Non-parametric estimation of the probability density function allows a more general description of the region distribution, since no assumption on the pdf is done. The simplest way to have a non-parametric estimation of the gray value distribution over a region is the histogram [31]. This method consists just on counting the

number of time that a pixel belongs to a bin or gray-level value. This method is fast but can give a sparse result if the number of sample is low. An alternative is to use kernel estimation which performs a local smoothing of the histogram. Then the pdf $p(I)$ associated with an observation I for a fixed region Ω at a given moment can be defined by the Parzen model [77] (case where the kernel is gaussian) as follows:

$$p(I) = \frac{1}{Z} \int_{\Omega} K(I - I(x)) dx, \quad (2.2)$$

where $|\cdot|$ is the area of the given region and $K(\cdot)$ is the 1- D Gaussian kernel with 0-mean and variance σ^2 . This estimation has been used in the segmentation framework in [49, 70, 85] and is the one that will be used in the second part of this thesis. In fact our framework must stay general and adjustable to any image type.

2.3 Conclusion

In this chapter a short review on extraction relevant features from textures was presented. Structure tensor and Gabor filters are probably the most used in the image segmentation literature. Besides of the problem of extracting features, one must consider also the possible measures to study these features.

The next chapter presents our new texture descriptor. We will compare it to the Gabor filter responses and the structure tensor.

A New Intrinsic Geometric Texture Descriptor based on the Principal Curvatures

3

In this chapter, we present a new texture descriptor based on the shape operator defined in differential geometry.

3.1 Textures and Beltrami Framework

In a general way, textures are difficult to define and no precise mathematical definition has been found so far. However, it is consensually admitted that textures are fine scale-details, usually with some periodicity and oscillatory nature [4]. Besides, textures raises the problem of non-existence of significant edges and the non-homogeneity of intensity distributions lying in images. The difficulty of having a mathematical definition for textures has induced different choice of texture representation. In this work, we are particularly interested in the Beltrami representation introduced by Sochen, Kimmel and Malladi in [94]. Sochen et al proposed a new efficient representation of images by considering images as a Riemannian manifold embedded in a higher dimensional space. For instance, a standard 2 dimensional gray value image $I : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ can be viewed as a surface Σ with local coordinates (x, y) embedded in \mathbb{R}^3 by a smooth mapping $X : (x, y) \rightarrow (X_1 = x, X_2 = y, X_3 = I(x, y))$ (Figure 3.1). This manifold-based representation of images offers two main advantages. Firstly, it allows to use efficient tools borrowed from differential geometry to perform different image processing tasks such as denoising or segmentation as we will do in the next chapters. The second main advantage is the ability to work with arbitrary N dimensional images. For example, a color image can also be represented in a 5 dimensional space by the mapping $X : (x, y) \rightarrow (X_1 = x, X_2 = y, X_3 = R(x, y), X_4 = G(x, y), X_5 = B(x, y))$, where R,G,B stands for red, green and blue. Sagiv, Sochen and Zeevi in [87] used the Beltrami framework to represent the texture image as a 2-D dimensional manifold embedded in a space of $N + 2$ dimensions, where N is the number of Gabor responses. They

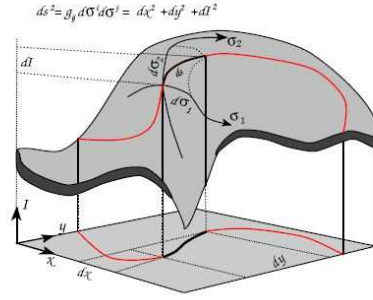


Figure 3.1: Images represented by Riemannian manifold embedded in a higher dimensional space. Reproduced from [94]

used the first fundamental form [57], also called metric tensor, of the texture manifold to define an intrinsic edge detector like in [89]. The idea of using the metric tensor to intrinsically define the edges between different texture regions is efficient in the context of differential geometry. Indeed, the first fundamental form describes the distortion or rate of change of the manifold and so can detect boundary between different parts of the manifold corresponding to different homogeneous textures. More precisely Sagiv et al used the geodesic active contour model [20] to drive the evolving contour toward the boundaries between two different texture regions by considering the edge detector function or stopping function as the inverse of the determinant of the metric tensor. This can be explained in the following way. If we consider the definition of the first fundamental form:

$$g_{\mu\nu} = \left(\left\langle \frac{\partial X}{\partial \mu}, \frac{\partial X}{\partial \nu} \right\rangle \right),$$

where $\mu, \nu = x, y$ in the (x, y) -basis. We have in the case of gray scale images, $X := (x, y, I)$ and

$$\begin{aligned} g_{xx} &= 1 + I_x^2, \\ g_{xy} &= I_x I_y, \\ g_{yy} &= 1 + I_y^2, \end{aligned}$$

which implies that

$$\frac{1}{\det(g_{\mu\nu})} = \frac{1}{(1 + |\nabla I|^2)}. \quad (3.1)$$

Function 3.1 corresponds exactly to the edge detector function used in the standard model [20]. Thus, Sagiv et al used the metric tensor of texture images to define an efficient edge detector for textural images. Nevertheless, as we said earlier, the edge detector function is not robust enough to segment a wide range of images and a region-based term, coming from the vectorial Chan-Vese model [23], was coupled with their intrinsic edge detector function to perform the segmentation of complex textures. This coupling is necessary because edge-based active contours are too sensitive to noise, bad contrast and initial position.

3.2 New Texture Descriptor based on the Beltrami Framework and the Shape Operator

Our goal is to define a region descriptor/feature, instead of a boundary descriptor, for textural regions. Like [87], we also wish to develop an intrinsic descriptor based on the geometrical shape of the manifold representing the texture region. We thus propose to use the *shape operator* and more specifically the eigenvalues of the shape operator to describe the geometry of the textures of interest. The shape operator is a linear operator which calculates the bending of a surface in different directions [45]. The eigenvalues of the shape operator correspond to the extremal of bending of the surface, they are called principal curvatures and they are known to represent the geometry of the considered smooth manifold. Indeed, in the simple case of a curve in the space, the curvature κ of this curve is the inverse of the radius ρ of the best fitting circle to the curve, i.e $\kappa = \frac{1}{\rho}$. It is then intuitive to say that the curvature κ describes the local shape of the curve and by extension the principal curvatures describe the manifold.

In this work, we choose to represent the texture manifold by the simplest Beltrami representation, i.e. $X = (x, y, I)$. Since texture images are seen as a 2-D manifold, two principal curvatures can be computed in this representation, namely (κ_1, κ_2) . The couple of principal curvatures (κ_1, κ_2) defines an intrinsic and efficient descriptor to segment complex textural regions. More precisely, we believe that for a given texture pattern, a distribution/pdf of couples (κ_1, κ_2) is repeated inside the texture region. This distribution will be automatically estimated through the segmentation process.

Let us introduce the mathematical definition of the shape operator that we call S . The shape operator measures the shape of the manifold in a given region by estimating how the normal \mathcal{N}_Σ to the surface Σ changes from point to point.

Definition 1. [45] *Let Σ be a regular surface, and let \mathcal{N}_Σ be a surface normal to Σ defined in a neighborhood of a point $\mathbf{p} \in \Sigma$. For a tangent vector $\mathbf{v}_\mathbf{p}$ to Σ at p , the shape operator is defined as:*

$$S(\mathbf{v}_\mathbf{p}) = -\mathbf{D}_{\mathbf{v}_\mathbf{p}}\mathcal{N}_\Sigma, \quad (3.2)$$

where $\mathbf{D}_{\mathbf{v}_\mathbf{p}}\mathcal{N}_\Sigma$ is the derivative of the surface normal \mathcal{N}_Σ in direction $\mathbf{v}_\mathbf{p}$.

Definition 2. [45] *The eigenvalues of the shape operator S of a regular surface Σ at $\mathbf{p} \in \Sigma$ are precisely the principal curvature of Σ at \mathbf{p} . The corresponding unit eigenvectors are unit principal vectors, and vice versa.*

In our situation:

Lemma 1. [57] *The principal curvatures κ_1, κ_2 of the 2-D manifold are the roots of the following equation:*

$$\kappa^2 - b_{\mu\nu}g^{\mu\nu}\kappa + \frac{b}{g} = 0, \quad (3.3)$$

where $g^{\mu\nu}$ is the inverse metric of $g_{\mu\nu}$, g, h are the determinant of $g_{\mu\nu}, b_{\mu\nu}$ and $b_{\mu\nu}$ is the second fundamental form defined by:

$$b_{\mu\nu} = \left(\left\langle \frac{\partial^2 X}{\partial \mu \partial \nu}, \mathcal{N}_\Sigma \right\rangle \right),$$

where $\mu, \nu = x, y$ in the (x, y) -basis, and the Einstein summation convention is used in (3.3), which means that elements with identical subscripts and superscripts are summed over.

Let I be the original gray level image. The mapping X is equal to $(x, y) \rightarrow (x, y, I(x, y))$ and the first fundamental form is thus given by

$$g_{\mu\nu} = \begin{pmatrix} 1 + I_x^2 & I_x I_y \\ I_x I_y & 1 + I_y^2 \end{pmatrix},$$

where the suffixes stands for partial derivatives. The normal to the manifold is given by

$$\mathcal{N}_\Sigma = \frac{1}{Z}(-I_x, -I_y, 1),$$

with $Z = \sqrt{1 + I_x^2 + I_y^2}$, which yields us to the calculus of the second fundamental form

$$b_{\mu\nu} = \frac{1}{Z} \begin{pmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{pmatrix}.$$

Using Lemma 1, the values of the principal curvatures are given by

$$\kappa_{1,2} = \frac{(-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma})}{2\alpha},$$

where

$$\begin{cases} \alpha = 1 \\ -\beta = \frac{1}{Z^3} [I_{xx}(1 + I_y^2) + I_{yy}(1 + I_x^2) - 2I_{xy}(I_x I_y)] \\ \gamma = [I_{xx}I_{yy} - (I_{xy})^2] \end{cases}.$$

The first principal curvature κ_1 ($\kappa_1 \geq \kappa_2$) corresponds to the maximal change of the normal to the surface and κ_2 corresponds to the minimum change. For sake of simplicity, and in order to use the information provided by the two principal curvatures, we consider to work with the norm of $\mathbf{k}_1 + \mathbf{k}_2$, where vector \mathbf{k}_1 (resp. \mathbf{k}_2) has a norm κ_1 (resp. κ_2) and is oriented by the associated unit principal vector (see Definition 2). Since \mathbf{k}_1 and \mathbf{k}_2 are orthogonal, this leads to:

$$\kappa_t := \sqrt{\kappa_1^2 + \kappa_2^2}, \quad (3.4)$$

where $\kappa_t : \Omega_0 \rightarrow \mathbb{R}_+$ defines the texture descriptor that we will use to segment regions with different texture patterns and Ω_0 corresponds to the image domain.

3.3 Results

In this section we discuss the results of our texture descriptor algorithm compared with Gabor filter response and structure tensor outputs.

Natural Textured Image We propose to look at a natural picture (Figure 3.2(a)) coming from the Berkeley data set. Figure 3.2(b) shows the result of our texture feature descriptor, where the feline is more discernable from the background. In fact, almost all the background has been set to a constant value. Figure 3.2(c) shows the Gabor filter response to the picture. It is difficult to interpret the quality of the feature extraction. However one can see that the information can be redundant or useless. Of course some channels seems to be good descriptors. Segmentation results based on these features will be shown in the next part.

Non-Oriented Texture The case where the textures that must be discriminated do not have a particular orientation is now studied. From a binary image 3.3(a) and from Brodatz Texture [15], a bi-textured image is constructed (Figure 3.3(b)). A plot of the probability density function on the object and the background is shown on Figure 3.3(c). The estimation has been done with Parzen window method (Equation 2.2). It is clear that the two density are too close to be discriminated by the distribution over their intensity value. The elements of the structure tensor (Eq. 2.1) are computed and shown on Figures 3.4(a) 3.4(b) 3.4(c) and the corresponding pdf Figures 3.4(d) 3.4(e) 3.4(f). As one can see the difference between the inside and the outside is not expressed by the distribution of structure tensor elements. Figure 3.5 shows the Gabor filter response for our bi-texture non-oriented image. Visually, none of the Gabor response seems enough discriminative, however we will see in Chapter 5 Section 5.4 that the combination of several of these channels gives interesting segmentation results. Finally we apply our texture descriptor in the non-oriented bi-texture image (Figure 3.6(a)). The object is already more easily distinguished from its background. The plots on Figure 3.6(b) of the object and background distribution upholds that our texture descriptor improves the discrimination between two different textures.

Oriented Texture We will finish by showing the limitation of our texture descriptor. Figure 3.7(a) presents the first test image composed of a vertical striped background with an horizontal striped object. In fact, the object is a background cut rotated of 90° . The degree of the curvature will not be different for a same uniformly oriented texture with different direction. Conversely this is the case where the Gabor filter is particularly efficient. Obviously, in this particular case, the selected Gabor responses corresponds to a orientation parameter $\theta = 0$ and $\theta = \frac{\pi}{2}$.

In summary, Gabor filters are powerful descriptors however they are computationally expensive. Moreover the space spanned by these features is high dimensional which will lead to a more complex segmentation framework. Structure tensor are interesting texture features when non-linear regularization is performed.

3.4 Conclusion

In this part we use a new texture descriptor based on the intrinsic local geometry representation of the image. We have been using the representation of the image as a surface in the image space as proposed by [93]. This representation offers the possibility to use geo-

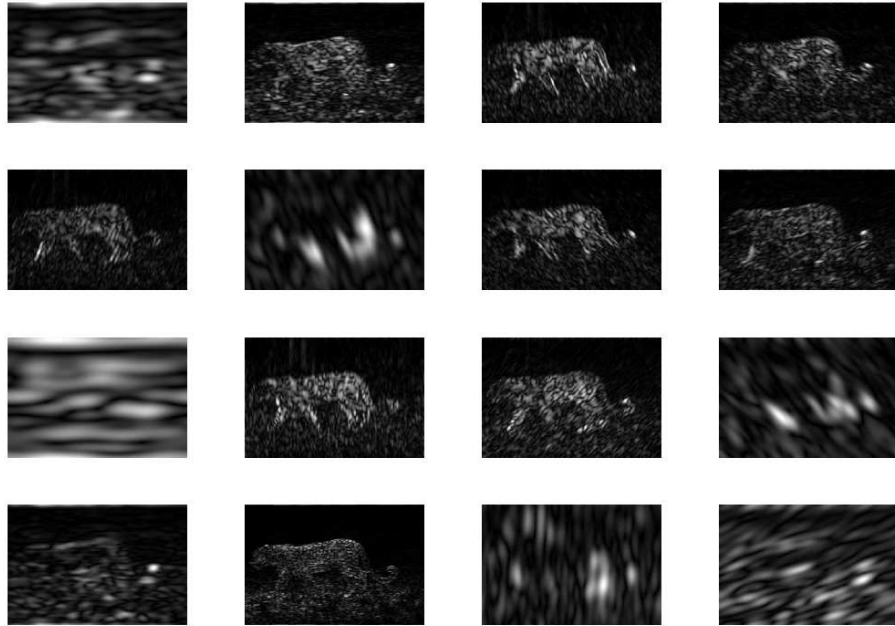
metric descriptors. In fact, from differential geometry, a new geometric texture descriptor, based on the shape operator, looks to be a promising texture feature to segment complex textures with different orientations and scales. From a numerical point of view, this texture descriptor is very easy to implement and an image of size 256×256 is processed in around 4 seconds. Our model is currently designed to work only with textures. Indeed, our model needs at least one textural region. If we consider a piecewise constant image then the value of our textural feature κ_t would be the same, equal to zero, on the whole image domain. Natural images are often composed of textured and piece-wise smooth regions and our feature descriptor is efficient to discriminate the textural part from the smooth part. We will see in the next part the importance of the texture feature for the segmentation task. In fact our texture descriptor based on the principal curvatures will be included in the segmentation framework of Chapter 5.



(a)



(b)



(c)

Figure 3.2: Natural images are often composed of textured and piece-wise smooth regions. (a) Natural Image. (b) Result of our texture feature based on the principal curvatures. (c) Gabor filter responses.

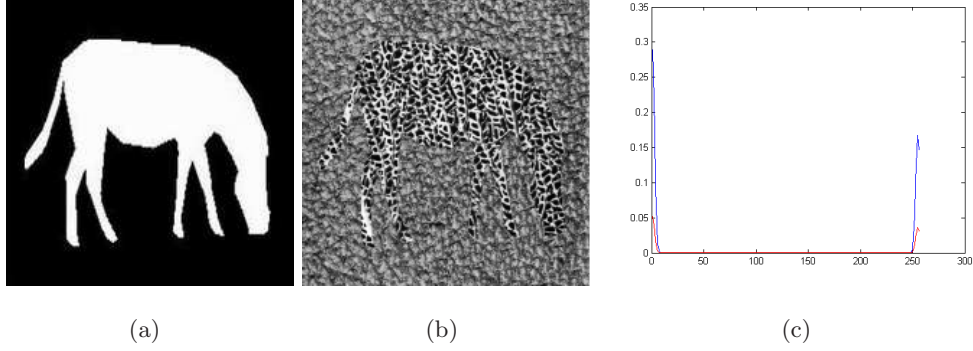


Figure 3.3: (a) Binary mask. (b) Original Texture Image constructed from Brodatz Texture [15]. (c) pdf on the object (red) and the background (blue).

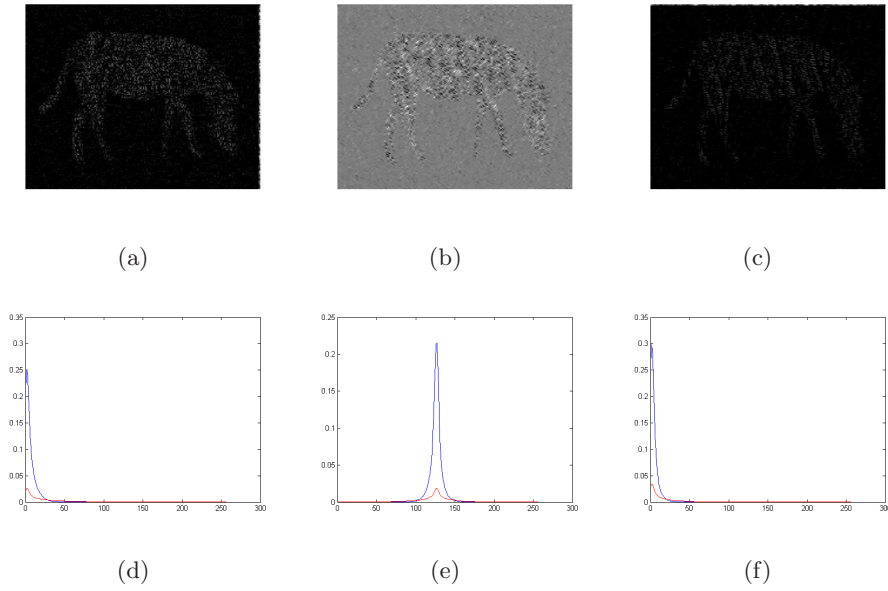


Figure 3.4: First row: Structure tensor elements on the texture of Figure 3.3(b), (a) I_x^2 , (b) $I_x I_y$, (c) I_y^2 . Second row: Correspondent pdf on the object (red) and the background (blue)

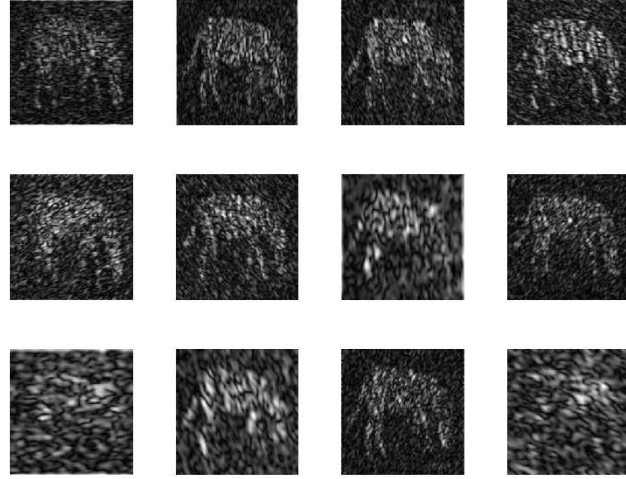


Figure 3.5: Gabor filter responses of figure 3.3(b)

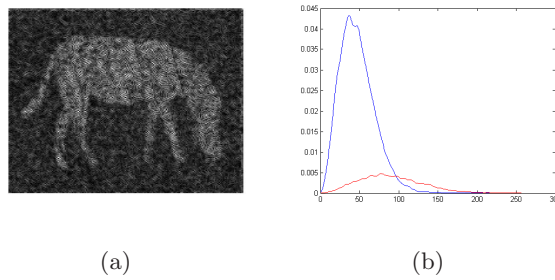


Figure 3.6: (a) Feature extracted from Figure 3.3(b) with our intrinsic feature descriptor.
(b) Correspondent pdf on the object (red) and the background (blue)

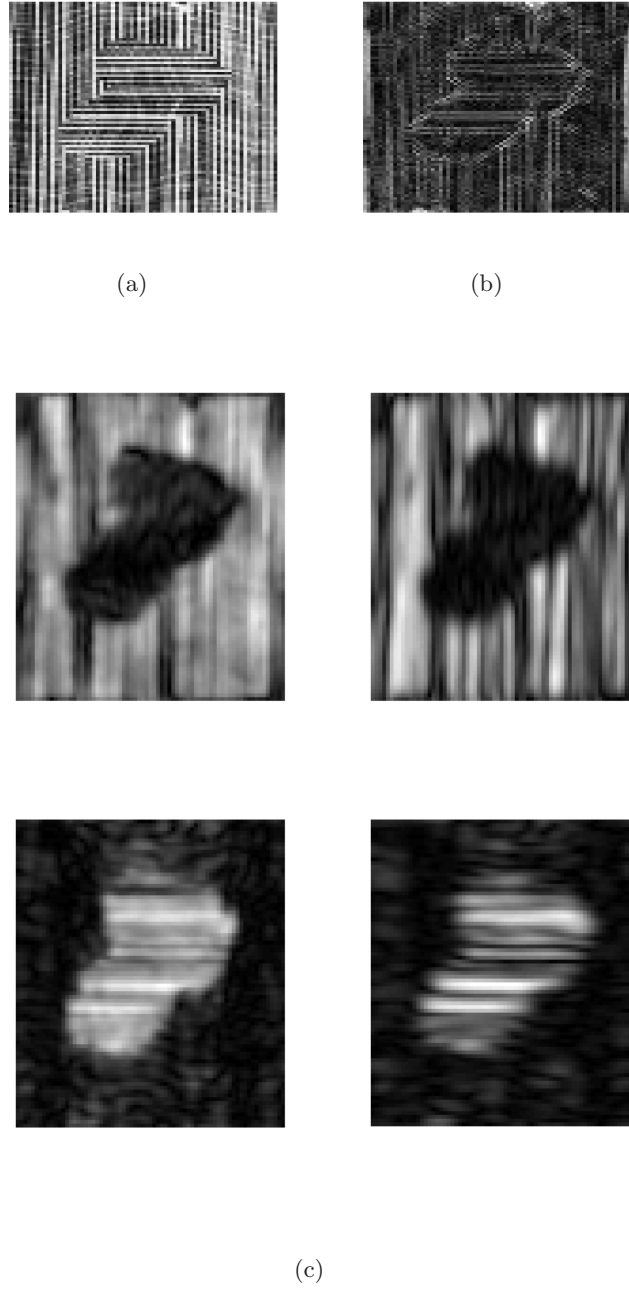


Figure 3.7: Failure example:(a)Textured image to be segmented. The background consists of vertical patterns and the object of horizontal patterns. (b)Feature extracted with our intrinsic feature descriptor (c) Features extracted with Gabor Filter

Part II

Unsupervised Segmentation

Motivation and Contribution

Motivation In the second part of this thesis an unsupervised segmentation algorithm is developed. The method is based on active contours (AC) and information theory.

Furthermore, we will also develop a fast and easy to implement algorithm to drive the active contour toward the minimum of the proposed energy functional. The first proposed model will be based on the popular Kullback-Leibler (KL) divergence, which will measure the difference between two probability density functions (pdf) of an image feature. We will maximize the KL using the algorithm introduced by Bresson et al in [13] in order to find distinct textured regions. A fast algorithm based on the dual formulation of the Total Variation (TV) norm proposed by Chambolle in [21] will be used to compute the global minimizer of the active contour model based on the KL distance. This model is constructed for 2 region segmentation. The second model deals with multi-region segmentation.

Contribution The main contributions of this part are summarized as follows:

- Definition of a new region-based energy using Kullback- Leibler divergence between inside and outside pdf of the evolving active contour. This technique can be viewed as a "probability density function competition".
- Analysis of the existence of a minimizing solution for that variational segmentation problem.
- A fast numerical scheme to determine the solution.
- A convex formulation of the multi-region active contours segmentation solved with a fast numerical scheme.

Background: Active Contours and Information Theory

4

4.1 Active contours

A lot of research has been going on in the field of image segmentation and many different segmentation methods have been developed over the last twenty years. Among these various models, the ones that are independent of any parametrization and that are based on a variational approach are our main interest. The review given in this section will cover the basics of variational models using boundary-based active contours [55], geodesic active contours [20], level set functions [73] and region-based segmentation [82]. The book of G.Aubert and P.Kornprobst [2] presents and explains most of variational methods for image segmentation. Some of them are summarized in the next section.

4.1.1 Boundary-Based Active Contours

Geodesic Active Contours and Variational Equations Active contours or snakes have been introduced by Kass et al. in 1987 [55]. The basic idea behind the active contour model is the following: the segmentation of any object in a given image $I \in L^1(\Omega)$ which is well discernible and whose edges can be described by a closed curve is equivalent to the location of sharp image intensity variations by iteratively deforming a curve C towards the edges of the object. The curve C has to be initialized close to the object of interest. Such a model is entirely dependent on the chosen parametrization of the initial curve $C(p) = (x(p), y(p)) \in \Omega$, $p \in [0, 1]$. The evolution equation of the curve $C(p)$ can be obtained by minimizing the following energy functional

$$F(C) = \alpha \int_0^1 \left| \frac{\partial C(p)}{\partial p} \right|_2 dp + \beta \int_0^1 \left| \frac{\partial^2 C(p)}{\partial p^2} \right|_2 dp + \lambda \int_0^1 g^2(I(C)) dp, \quad (4.1)$$

where $|\cdot|_2$ is L^2 -norm, α, β, λ are positive constants and g is an edge detecting function. The first two terms in Equation (4.1) are called *internal energy* and set up constraints on

the geometry of the active contour, i.e impose smoothness on the contour.

The third term is referred to as *external energy* and serves to attract the curve $C(p)$ toward the object's boundaries. One possible expression of g is:

$$g(I) = \frac{1}{1 + \gamma |\nabla(I * G_\sigma)|_2},$$

where γ is an arbitrary positive constant, G_σ is the Gaussian function with standard deviation σ and $(I * G_\sigma)$ represents a smoothed version of the original image I .

Besides the already mentioned dependency of the final segmentation result on the chosen parametrization, another disadvantage of the active contour model, such as initially proposed by Kass et al. [55], is its inability to handle topological changes.

Based on the method proposed in [55], Caselles et al. and Kichenassamy et al. in [20, 56] proposed the geometrically intrinsic model of geodesic/geometric active contours where the evolution curve is handled by the level set method introduced by Osher and Sethian in [73]. The problem of parametrization dependency is then solved. The new energy functional, which is independent of the initial curve parametrization, is given by:

$$F(C) = \int_0^1 g(|\nabla I_0(C(p))|) |C'_p| dp = \int_0^{L(C)} g(|\nabla I_0(C(s))|) ds, \quad (4.2)$$

where ds is the Euclidean element of length and $L(C)$ is the Euclidean length of the curve C , which is defined by $L(C) = \int_0^1 |C'_p| dp = \int_0^{L(C)} ds$. This energy functional can be seen as a weighted length of the curve C . As proved by Caselles et al. in [20], the minimization of the Energy Functional (4.2) results in a curve which is a geodesic in a Riemannian space. Minimizing Functional (4.2) is equivalent to minimize Functional (4.1) under the assumption that $\beta = 0$. This property has been shown by Caselles et al using concepts of Hamiltonian theory. Aubert and Blanc-Féraud [1] proved also this equivalence using calculus of variation which is a more natural tool for image processing community.

The minimization problem of any energy functional of form $F(c) = \int_a^b f(c, c_p) dp$ can be solved by using the calculus of variations as

$$\left[\frac{\partial}{\partial c} - \frac{d}{dp} \frac{\partial}{\partial c_p} \right] f(c, c_p) = 0. \quad (4.3)$$

Equation (4.3) is the well known Euler-Lagrange differential equation and any function c satisfying this equation corresponds to an extremum of $F(c)$.

The next step is the use of Equation (4.3) to minimize the geodesic active contour Energy Functional (4.2), where $f(C, C_p) = g(I(C)) |C_p|$ and thus Equation (4.3) becomes:

$$\left[\frac{\partial}{\partial C} - \frac{d}{dp} \frac{\partial}{\partial C_p} \right] (g(C) |C_p|) = 0. \quad (4.4)$$

Finding the evolution equation for geodesic active contours leads to solve the Euler-Lagrange Equation (4.4) by a gradient descent scheme. Let t be an artificial time step. The gradient descent formula which yields the minimum of the Euler-Lagrange function F , i.e. that can be used to find the C for which $\frac{\partial F}{\partial C} = 0$, then becomes

$$\frac{dC}{dt} = -\frac{\partial F}{\partial C}, \quad (4.5)$$

where the sign inversion in the right hand side is required in order to find a minimum and not a maximum. Using Equations (4.4) and (4.5), the final evolution equation for geodesic active contours is:

$$\frac{\partial C}{\partial t} = (g\kappa - \langle \nabla g, \mathcal{N} \rangle) \mathcal{N}. \quad (4.6)$$

The first term of the right hand side of Equation (4.6) represents the mean curvature motion weighted by the edge detecting function g . The curvature κ measures how fast the curve bends at any point and it is responsible for the evolution of the geodesic active contour in regions without edges and where the edge detecting function g approaches one, such as it is the case inside the object of interest. But more importantly, the curvature κ makes the geodesic active contour to become smoother by decreasing its total length. The second term of the equation's right hand side attracts the curve to the object's boundaries.

In the general case, the active contour is driven by the following flow:

$$\frac{\partial C}{\partial t} = V \mathcal{N}.$$

Level Sets Level sets functions are a non-parametric representation of curves. As already mentioned, level sets were first introduced by Osher and Sethian in [73] and they are a powerful way to handle topology changes of an evolving contour. The basic idea is to take the original interface and to add an extra dimension to the problem. In a two dimensional scenario, the original curve would lie in the xy plane at level zero ($z = 0$) and the z direction would then be used to measure the height, just as a topographic map indicating surface elevations. The interface lying at level zero is called 'zero level set' and represents the best segmentation result of the current iteration. Therefore, the goal is to let the level set function evolve in time and within a fixed coordinate system such that at each time instant t the outline of the curve's cross-section at height zero corresponds to the segmentation's evolving contour. By convention, the level set function is negative ($z < 0$) for all points whose x and y coordinates fall inside the zero level set and positive otherwise. Since the level set function is able to divide at any arbitrary height z , thereby taking over the shape of a mountain with two peaks, as illustrated in Figure 4.1, it becomes clear that by moving the level set function downwards, the zero level set is shifted from the base of the inverted mountain up to the two peaks and the initially single closed curve at level zero splits in order to form two closed curves at the new location of the level zero. Starting with a single evolving contour, level sets make it therefore possible to segment two separate objects.

Being familiar with the general idea behind the level set method, it is now interesting to analyze how the level set function actually evolves. The general evolution equation for a two-dimensional curve is given by the partial differential equation

$$\begin{cases} \frac{dC}{dt} = V_{\parallel} \mathcal{T} + V_{\perp} \mathcal{N}, \\ C_{t=0} = C_0, \end{cases} \quad (4.7)$$

where T and N are respectively the unit tangential and unit normal to the curve C and V_{\parallel} and V_{\perp} are the tangential and normal velocities of the evolving front. Epstein and Gage [35] have shown that the tangential velocity component V_{\parallel} has no effect on the deformation of the curve and can therefore be neglected. Equation (4.7) holds for any hyper-surface which is represented either parametrically or implicitly, as in the case of level sets. The first step

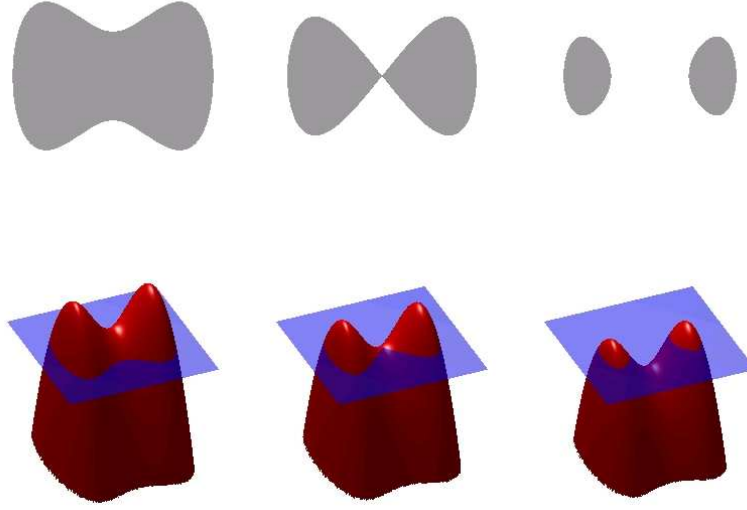


Figure 4.1: Evolution of a level set function illustrating its ability to handle topology changes. Source: Wikipedia.

in the derivation of the evolution equation of level sets is to create an initial time dependent level set function $\phi(\mathbf{x}, t = 0)$ based on the signed distance d from each point $\mathbf{x} = (x, y)$ to an initial closed front $\Gamma(t = 0)$. The level set function is thus initialized in our case with a distance map having negative values for points lying inside the front and positive values for all points lying outside and it satisfies the following definitions:

$$\begin{cases} \phi(\mathbf{x}, t) < 0 & \text{for } \mathbf{x} \in \Omega_{in}(t) \\ \phi(\mathbf{x}, t) > 0 & \text{for } \mathbf{x} \in \Omega_{out}(t) \\ \phi(\mathbf{x}, t) = 0 & \text{for } \mathbf{x} \in \Gamma(t), \end{cases}$$

where Ω_{in} and Ω_{out} are the region inside and outside the contour $\Gamma(t)$. Given that the evolving front corresponds to the zero level set, meaning to the front where $\phi(x(t), y(t), t) = 0$, the following equality must be observed:

$$\phi(x(t), y(t), t) = 0 \quad \forall t.$$

Thus the evolution of the front Γ , moving in the normal direction to itself with the speed V_{\perp} , can be described by means of the so-called Hamilton-Jacobi equation applied to the level set function:

$$\frac{\partial \phi}{\partial t} + V_{\perp} \|\nabla \phi\| = 0. \quad (4.8)$$

In the last step, we are going to apply the level set evolution Equation (4.8) to geodesic active contours. By comparing Equation (4.6) with Equation (4.7) we can immediately identify the normal speed V_{\perp} :

$$\frac{\partial C}{\partial t} = \underbrace{(g\kappa - \langle \nabla g, \mathcal{N} \rangle)}_{V_{\perp}} \mathcal{N}. \quad (4.9)$$

By substituting the found velocity V_\perp into the level set evolution Equation (4.8) and by knowing that the curvature κ and the unit normal to the evolving front \mathcal{N} are respectively given by $\kappa = \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right)$ ([90]) and $\mathcal{N} = \frac{\nabla \phi}{|\nabla \phi|}$, we finally get the wanted evolution equation of the level set function ϕ

$$\phi_t = \{g\kappa - \langle \nabla g, \mathcal{N} \rangle\} |\nabla \phi| \quad (4.10)$$

Finally, we note that the active contour and the level set formulation are equivalent:

$$\frac{\partial C}{\partial t} = V\mathcal{N} \Leftrightarrow \frac{\partial \Phi}{\partial t} = V|\nabla \phi|.$$

4.1.2 Region-Based Segmentation

Originally, active contours, such as introduced by Kass et al. in [55], were boundary based methods, meaning that the earliest evolution equations included only terms based on local information on the boundaries of the object of interest. The same holds for the geodesic active contours model, that have been proposed by Caselles et al. in [20] and Kichenassamy in [56]. Later on, more sophisticated segmentation models emerged that integrated also global information on the regions of interest additionally to the information based on object boundaries. Pioneers in this respect were Cohen et al. [26] and Ronfard [82]. Cohen et al. presented in [26] a surface reconstruction method using region based active contours. Later, it was Ronfard who proposed in [82] that an evolution function of region based active contours should be proportional to the difference of statistical features: $E = k^{(in)} - k^{(out)}$, where $k^{(in)}$ and $k^{(out)}$ represent statistical models of the regions inside and outside the object of interest, respectively. Thereafter, a statistical framework for image segmentation has been presented by Zhu and Yuille [110], which is known under the name of *region competition*. In this model, the active contour evolution equation was obtained by minimizing a generalized Bayes criterion based on the Mumford-Shah functional. The method of Zhu and Yuille has been extended by Paragios and Deriche [74], who improved the contour descriptor in order to incorporate the image gradient, such as in geodesic active contours, and the evolution equation is handled by level set method.

Further, Chan and Vese proposed in [22] a 2-phase segmentation method based on the mean descriptor where the active contour evolves in such a way that the difference between the inside (resp. outside) gray level value and the inside (resp. outside) mean value c_{in} (resp. c_{out}) is minimized. This can be expressed by the following energy functional:

$$F(c_{in}, c_{out}, C) = \int_{\Omega_{in}} (I_0(x, y) - c_{in})^2 dx dy + \int_{\Omega_{out}} (I_0(x, y) - c_{out})^2 dx dy + \nu |C|, \quad (4.11)$$

where $|C|$ is the contour length and ν a positive weight parameter. We can note that Energy 4.11 is a particular case of the Mumford-Shah Functional [71].

Then, a general Eulerian framework for region based active contours has been proposed by Jehan-Besson and Barlaud in [53]. They used region dependent descriptors that were globally attached to the evolving regions and that allowed to simultaneously perform the segmentation of the object of interest together with the estimation of these region descriptors. The probability density function (pdf) looks so far to be one of the most efficient region descriptor to solve the segmentation problem. More recently, Aubert, Barlaud, Faugeras, Jehan-Besson and Herbulot in [3, 48, 53] proposed to update the pdf of the object and the

background during the segmentation object until the optimal partition is reached. This idea makes a lot of sense because segmentation and estimation of features such as pdfs of regions of interest are basically related to each other. Besides, this approach does not need a pre-processing step to estimate the optimal features. We will apply this idea in our proposed segmentation method, combined with tools coming from information theory presented in Section 4.3.

4.2 Global Minimization of Active Contour Energy

The two principal drawbacks of active contours methods are the computational cost and the existence of local minimizers. In most papers regarding active contours, the Euler-Lagrange equation of the variational model is discretized using an explicit scheme, which produces a slow segmentation process. Besides, most active contour evolutions are handled by the level set method [73], which needs to use a signed distance function re-computed regularly during the evolution process to avoid numerical instabilities. Moreover the problem of local minimizers must be underlined. A segmentation problem based on variational model where optimizers are local implies that the solution is highly dependent of the initial condition. A recent method, introduced by Bresson et al. [13], proposes to redefine the active contour model into a model which gives global minimizers. Furthermore a fast numerical scheme can solve the segmentation problem. Thus the global minimization of the AC energy can be derived in two different ways, either from the Euler-Lagrange equation or from the variational model.

4.2.1 Model 1: Global Active Contours from Partial Differential Equation

The pioneer work of Chan, Esedoglu and Nikolova in [24] unifies image denoising and image segmentation through the Chan-Vese model [22]. Global minimizers of the segmentation and denoising model are then defined. Based on [24], Bresson et al. in [13], proposed to compute a global minimum to the active contour energy including the case of geodesic active contour [20] and segmentation based on the general Mumford-Shah energy, in order to be independent of the initial contour position. Following [13], we explicit hereafter how the active contour model can be turned into a convex model in a general case.

Let us consider the following variational model:

$$\min_{\Omega_C} \left\{ F_1(\Omega_C) = \int_{\partial\Omega_C} g(s)ds + \lambda \left[\int_{\Omega_C} r^{in}(x)dx + \int_{\Omega \setminus \Omega_C} r^{out}(x)dx \right] \right\}, \quad (4.12)$$

where g is an edge detection function, Ω_C is the evolving region and $\Omega_C \cup \overline{\Omega_C} = \Omega$, r^{in} (resp. r^{out}) is a model of the inside (resp. outside) region.

This model merges the boundary and the region information of the image. The associated Euler-Lagrange Equation is given by:

$$\phi_t = \left(\nabla \frac{\nabla \phi}{|\nabla \phi|} + \lambda r \right) |\nabla \phi|, \quad (4.13)$$

where $r = r^{in} - r^{out}$. For instance the Chan-Vese mean descriptor gives

$$r = ((I(x) - c_{in})^2 - (I(x) - c_{out})^2),$$

where c_{in} is the mean inside the object and c_{out} the mean outside. Since $|\nabla\phi| > 0$, the steady state of (4.13) is the same as:

$$\frac{\partial\phi}{\partial\tau} = -\lambda r + \nabla \frac{\nabla\phi}{|\nabla\phi|}. \quad (4.14)$$

Then, ϕ in (4.14) is also solution of the variational model:

$$\min_{\phi \in [0,1]} \left\{ F(\phi) = \int_{\Omega} -\lambda r \phi + |\nabla\phi| \right\}. \quad (4.15)$$

By restricting the minimization of ϕ in the convex set of element in $[0, 1]$, a solution can be found to the minimization problem.

To avoid any confusion with the level set function ϕ , the notation is changed and we are seeking the minimum of the functional $F(u)$ such that:

$$\min_{u \in [0,1]} \left\{ F(u) = \int_{\Omega} -\lambda r u + \int_{\Omega} |\nabla u| \right\}, \quad (4.16)$$

where $\int_{\Omega} |g(x)\nabla u| =: TV_g(u)$ is the weighted total variation norm of the function u .

Let us define the *characteristic function* $\mathbf{1}_{\Omega_C}$ of the set Ω_C , defined as follows:

$$\mathbf{1}_{\Omega_C}(x) = \begin{cases} 1 & \text{if } x \in \Omega_C \\ 0 & \text{otherwise} \end{cases}.$$

Theorem 4.2.1. *Suppose that $r \in C^1(\Omega)$ and $\lambda \in \mathbb{R}_+$, if u_* is any minimizer of $F(\cdot)$, then for almost every $\mu \in [0, 1]$ we have that the characteristic function*

$$\mathbf{1}_{\Omega_C(\mu)=\{x:u(x)>\mu\}}(x),$$

where C is the boundary of the set Ω_C , is a **global minimizer** of $F(\cdot, \lambda)$.

Proof. See [13, 24]. \square

The variational model (4.16) can be quickly minimized using a dual approach of the TV norm as in [4, 13, 21]. A convex regularization of the variational model (4.16) is used:

$$\min_{v \in [0,1], u} \left\{ F(u, v) = \int_{\Omega} |\nabla u| - \lambda r v + \frac{1}{2\theta} (u - v)^2 \right\}, \quad (4.17)$$

where $\theta > 0$. Since the functional F is convex w.r.t. u, v , its minimizer can be computed by minimizing F w.r.t. u, v separately, and iterating until convergence as described in the references mentioned above. Thus, the following minimization problems are considered:

$$v \text{ being fixed, } \min_u \left\{ TV(u) + \frac{1}{2\theta} \|u - v\|_{L^2}^2 \right\}, \quad (4.18)$$

$$u \text{ being fixed, } \min_{v \in [0,1]} \left\{ \int_{\Omega} -\lambda r v + \frac{1}{2\theta} (u - v)^2 \right\}, \quad (4.19)$$

- The solution of (6.16) is given by [21]:

$$u = v - \theta \nabla \mathbf{p}, \quad (4.20)$$

where $\mathbf{p} = (p^1, p^2)$ is given by

$$\mathbf{p}^{n+1} = \frac{\mathbf{p}^n + \delta t \nabla (\operatorname{div} \mathbf{p}^n - v/\theta)}{1 + \frac{\delta t}{g(x)} |\nabla (\operatorname{div} \mathbf{p}^n - v/\theta)|}, \quad n \in \mathbb{N}$$

- and the solution of (6.17) is given by [13]:

$$v = \min\{\max\{u + \theta \lambda r, 0\}, 1\}. \quad (4.21)$$

4.2.2 Model 2: Global Active Contours from the Variational Model

Let us consider a slightly different version of Functional 4.23:

$$\min_{u \in [0,1]} \{F_2(u) = \int_{\Omega} g(x) |\nabla u(x)| dx + r^{in}(x) u(x) dx + r^{out}(x) (1 - u(x)) dx\} \quad (4.22)$$

This functional has been introduced by Mory and Ardon [69] as a fuzzy region competition for convex two-phase segmentation. In fact this function is convex in u , $u \in [0, 1]$ and the following property can be deduced:

Proposition 4.2.1. *If $u = \mathbf{1}_{\Omega_C}$ Energy from Functional 4.15 and Energy from Functional 4.22 are equivalent .*

$$F_2(u = \mathbf{1}_{\Omega_C}) = F_1(\Omega_C).$$

We can also notice that the minimization of Energy 4.22 is equivalent to the minimization of Energy 4.15. Then Theorem 4.2.1 remains valid for Energy 4.22 and the minimization scheme for fixed region descriptor is the same as for Model 1. This algorithm has shown its efficiency on medical image as in [69].

4.2.3 Model 1 vs Model 2

The two presented methods for global minimizers for active contour model are very similar. Model 1 was the first model to offer an equivalent convex formulation of the active contour model. Since the solution is given from the active contour general evolution equation, any active contour formulation can be adapted. The difference between the two models lies in the computation of the optimal region parameters. For instance let us consider the Chan-Vese model [22]:

Model 1

$$\begin{aligned} \min_{u, c_{in}, c_{out}} \{F_1(\Omega_C, c_{in}, c_{out}) &= \int_{\partial\Omega_C} g(s) ds + \int_{\Omega_C} ((I(x) - c_{in})^2 + \int_{\Omega \setminus \Omega_C} (I(x) - c_{out})^2) dx\}, \\ c_{in} &= \frac{\int_{\Omega} \mathbf{1}_{\{u > \mu\}}(x) I(x)}{\int_{\Omega} \mathbf{1}_{\{u > \mu\}} u(x) dx} \quad c_{out} = \frac{\int_{\Omega} \mathbf{1}_{\{u < \mu\}} I(x)}{\int_{\Omega} \mathbf{1}_{\{u < \mu\}} dx} \end{aligned}$$

where μ is an arbitrary real value in $[0, 1]$.

Model 2

$$\begin{aligned} \min_{\Omega_C, c_{in}, c_{out}} \{F_1(\Omega_C, c_{in}, c_{out}) &= \int_{\partial\Omega} g(s)|\nabla u|dx + \int_{\Omega} (I(x) - c_{in})^2 u(x) + \int_{\Omega} (I(x) - c_{out})^2 (1 - u(x))dx\}, \\ c_{in} &= \frac{\int_{\Omega} u(x)I(x)}{\int_{\Omega} u(x)dx} \quad c_{out} = \frac{\int_{\Omega} (1 - u(x))I(x)}{\int_{\Omega} (1 - u(x))dx} \end{aligned}$$

The estimation of the parameters by Model 2 is less rigid then by Model 1. Moreover there is no dependency to the arbitrary parameter μ . However the model supposes that the energy can be expressed as a difference between regions which is not always the case as we will see in Chapter 5.

4.2.4 Numerical Scheme

The minimization of Energy 4.16 and Energy 4.22 is based on the two iteration schemes 4.20 and 4.21. The numerical scheme to compute the divergence and the gradient operator for 3 dimensional images are given as follows: $\mathbf{p} = (p^1, p^2, p^3)$

$$\begin{aligned} (div(\mathbf{p}))_{x,y} = & \begin{cases} p_{x,y,z}^1 - p_{x-1,y,z}^1 & \text{if } 1 < x < N_x, \\ p_{x,y,z}^1 & \text{if } x = N_x, \\ -p_{x-1,y,z}^1 & \text{if } x = N_x, \end{cases} + \begin{cases} p_{x,y,z}^2 - p_{x-1,y,z}^2 & \text{if } 1 < y < N_y, \\ p_{x,y,z}^2 & \text{if } y = N_y, \\ -p_{x,y-1,z}^2 & \text{if } y = N_y \end{cases} \\ & + \begin{cases} p_{x,y,z}^2 - p_{x,y,z-1}^2 & \text{if } 1 < z < N_z, \\ p_{x,y,z}^2 & \text{if } z = N_z, \\ -p_{x,y,z-1}^2 & \text{if } z = N_z \end{cases} \end{aligned}$$

The discrete gradient operator is as follows:

$$\begin{aligned} (\nabla u)_{x,y,z}^1 &= \begin{cases} u_{x+1,y,z} - u_{x,y,z} & \text{if } x < N_x, \\ 0 & \text{if } x = N_x \end{cases} \\ (\nabla u)_{x,y,z}^2 &= \begin{cases} u_{x,y+1,z} - u_{x,y,z} & \text{if } y < N_y, \\ 0 & \text{if } y = N_y \end{cases} \\ (\nabla u)_{x,y,z}^3 &= \begin{cases} u_{x,y,z+1} - u_{x,y,z} & \text{if } z < N_z, \\ 0 & \text{if } z = N_z \end{cases} \end{aligned}$$

Deducing the divergence and the gradient operator for 2 dimensional images is straightforward. In [21], the authors shows that the algorithm converges if $\delta t < \frac{1}{8}$ for 2D images and we can easily see from his proof that for 3D images the convergence is guarantied for $\delta t < \frac{1}{16}$. Thus for algorithms proposed in the next chapter δt is fixed to the value $\frac{1}{8}$ for 2D images and $\frac{1}{16}$ for 3D images.

4.3 Information Theoretic and Probabilistic tools

Information theory has provided powerful tools which have been successfully applied to solve various problems in science and technology. We recall hereafter some general definitions on probability and information theory that will be useful to define an image segmentation model.

4.3.1 Shannon Entropy

Information entropy has been introduced by Claude Shannon in 1948 in [91]. Shannon's definition of entropy $H(A)$ of any continuous random variable A is given by:

$$H(A) = - \int_{a \in \Omega_A} p(a) \log p(a) da,$$

where Ω_A is the set of possible outcomes of A and $p(\cdot)$ is the probability measure of A .

4.3.2 Relative Entropy

Let A be a continuous random variable defined on the set of possible outcomes Ω_A and $p(A)$ and $q(A)$ two probability measures of A . The relative entropy is a measure of the distance between the probability density functions p and q defined as follow:

$$KL(p, q) = \int_{a \in \Omega_A} p(a) \log \frac{p(a)}{q(a)} da.$$

Since the relative entropy is not a true metric (it does not satisfy the symmetric property of distance function), an extension called the symmetric Kullback-Leibler distance has been widely used. The definition of the Kullback-Leibler (KL) distance is thus as follows:

$$KL(p, q) = \int_{a \in \Omega_A} \left\{ p(a) \log \frac{p(a)}{q(a)} + q(a) \log \frac{q(a)}{p(a)} \right\} da. \quad (4.23)$$

The well-known mutual information is in fact a particular case of the KL distance, where the distance between the joint distribution and the product distribution is measured. Let $p(A_1)$ and $q(A_2)$ be two probability measures associated respectively to the continuous random variables A_1 and A_2 . The mutual information is defined as:

$$MI(p, q) = \int_{a_1 \in \Omega_A} \int_{a_2 \in \Omega_A} p(a_1, a_2) \log \frac{p(a_1, a_2)}{p(a_1)q(a_2)} da_1 da_2,$$

where $p(a_1, a_2)$ is the joint distribution computed from the joint histogram. Mutual Information is often defined as the mutual dependence of the two random variables A_1 and A_2 .

Information theory has been widely used for image processing in different type of tasks. For instance, mutual information is one of the most "popular" similarity measures for multimodal registration process (see [79] for a review). The entropy measurement is encountered for probabilistic image classification model as well for image segmentation. In this work, we propose to apply the KL's approach to the image segmentation task in an unsupervised way.

4.3.3 From Image Space to Probability Space

A given image is often modeled as a scalar function $I : \Omega \rightarrow (R)$, where Ω is a regular open bounded set of \mathbb{R}^N corresponding to the image domain. The image is a deterministic signal. The usual way to pass from the deterministic world to the probabilistic one is to use a probability density estimation. In this work, the non-parametric estimation of pdf called

the Parzen window density estimation is used. In the standard formulation from [77], the Parzen window method estimates the pdf from N samples $\gamma_1 \dots \gamma_N$ of a random variable Γ from the following equation:

$$p(\Gamma = \gamma) = \frac{1}{N} \sum_{i=1}^N K(\gamma - \gamma_i),$$

where K is the Gaussian kernel with 0-mean and variance σ^2 such that $K(\cdot) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-(\frac{(\cdot)^2}{2\sigma^2}))$ and γ describes the feature of interest.

For variational models, the image is defined in the continuous space, then a continuous version of the Parzen window estimation has to be considered. Hence, let Γ be a continuous random variable, if Γ depends on the spacial variable x defined on the image domain Ω then the estimated pdf is as follows:

$$p(\Gamma = \gamma) = \frac{1}{Z} \int_{x \in \Omega_0} K(\gamma - \gamma(x)) dx, \quad (4.24)$$

where the partition function Z is chosen to satisfy the probability axiom where $\int_{D_\Gamma} p(\gamma) d\gamma = 1$, D_Γ being the set of possible outcomes.

Typically, like in [50, 101], the random variable Γ is the image intensity and the set $D_\Gamma = \mathbb{R}$ is the histogram domain. In this case, the constant Z that satisfies the probability measure requirement $\int_{\mathbb{R}} p(\gamma) d\gamma = 1$ is given by $Z = \int_{\mathbb{R}} \int_{\Omega} K(I - I(x)) dx dI = |\Omega|$, where $|\Omega|$ is the area of Ω . The pdf is estimated by the parzen method:

$$p(I) = \frac{1}{|\Omega_0|} \int_{\Omega_0} K(I - I(x)) dx, \quad (4.25)$$

and the KL divergence measure between p and q is expressed by:

$$KL(p(I), q(I)) = \int_{\mathbb{R}} \left\{ p(I) \log \frac{p(I)}{q(I)} + q(I) \log \frac{q(I)}{p(I)} \right\} dI. \quad (4.26)$$

4.4 Conclusion

In this chapter, we introduced the concept of image segmentation by active contours, as well as some concepts in probability and information theory that will directly serve us. In the next chapter, we will introduce our second contribution: an unsupervised active contour segmentation based on the Kullback-Leiber distance.

Unsupervised Region Competition Based on Active Contour and Kullback-Leibler Distance

5

In this chapter we introduce our unsupervised segmentation model. Our model is a region-based active contour method. We will see that our energy holds good mathematical properties. First the existence of a minimizer is shown and second a fast minimization algorithm is defined.

5.1 Proposed 2-Phase Segmentation Method

Based on information theory (Section 4.3), we develop our unsupervised image segmentation model below.

We propose to carry the unsupervised image segmentation task for gray-scale images which are assumed to be composed of an object and a background. An efficient way to perform an unsupervised segmentation is to use the Region Competition approach as introduced by Zhu and Yuille in [110]. Here, a *pdf competition* approach is proposed based on the Kullback-Leibler (KL) divergence which measures a distance between two pdfs. We propose to maximize the KL distance between the probability density function inside and outside the active contour, which define two regions representing the object of interest and the background.

Let q_{in} be the inside probability density function, q_{out} the outside one, $\Omega = \Omega_{in}$ be the evolving region and $\Omega_0 \setminus \Omega = \Omega_{out}$ its complementary in the image domain Ω_0 . In this approach, the image intensity I is considered as a random variable. The set of possible outcomes is thus \mathbb{R} and the pdfs q_{in} and q_{out} associated with an observation I for a given

region Ω at a fixed moment are defined by:

$$\begin{cases} q_{in}(I, \Omega) &= \frac{1}{|\Omega|} \int_{\Omega} K(I - I(\hat{x})) d\hat{x}, \\ q_{out}(I, \Omega) &= \frac{1}{|\Omega_0 \setminus \Omega|} \int_{\Omega_0 \setminus \Omega} K(I - I(\hat{x})) d\hat{x} \end{cases}$$

where $|\cdot|$ is the area of the given region. The new (KL) distance is thus as follows:

$$\begin{aligned} KL(q_{in}(\Omega), q_{out}(\Omega)) &= \int_{\mathbb{R}} \left(q_{in}(I, \Omega) \log \left(\frac{q_{in}(I, \Omega)}{q_{out}(I, \Omega)} \right) + q_{out}(I, \Omega) \log \frac{q_{out}(I, \Omega)}{q_{in}(I, \Omega)} \right) dI \\ &= \int_{\mathbb{R}} (q_{in}(I, \Omega) - q_{out}(I, \Omega)) (\log q_{in}(I, \Omega) - \log q_{out}(I, \Omega)) dI. \end{aligned} \quad (5.1)$$

The Functional (5.1) gives a measure of difference between the pdfs defined inside and outside a given active contour represented in the region Ω . We naturally want to maximize Functional (5.1) in order to determine two regions with two pdfs as disjoint as possible, which provides two semantically different objects, expected to be the object of interest and the background. Maximizing Functional (5.1) involves the computation of its derivatives w.r.t the evolving domain Ω , which can be done with the shape derivative tool [3, 53] described in Annex A.1.

The Eulerian derivative in the direction \mathbf{V} of the criterion (5.1) is as follows:

$$\begin{aligned} \langle KL', \mathbf{V} \rangle &= \int_{\partial\Omega} \left\{ \frac{1}{|\Omega|} \int_{\mathbb{R}} \left(1 - \frac{q_{out}(I, \Omega)}{q_{in}(I, \Omega)} + \log \frac{q_{in}(I, \Omega)}{q_{out}(I, \Omega)} \right) \cdot [-K(I - I(s)) + q_{in}(I, \Omega)] dI \right. \\ &\quad \left. - \int_{\mathbb{R}} \frac{1}{|\Omega_0 \setminus \Omega|} \left(1 - \frac{q_{in}(I, \Omega)}{q_{out}(I, \Omega)} + \log \frac{q_{out}(I, \Omega)}{q_{in}(I, \Omega)} \right) \cdot [-K(I - I(s)) + q_{out}(I, \Omega)] dI \right\} \\ &\quad \cdot \langle \mathbf{V}(s), \mathcal{N}(s) \rangle ds, \end{aligned} \quad (5.2)$$

where \mathcal{N} is the unit inward normal to $\partial\Omega$ the boundary of the evolving region Ω , ds its length/area element. According to the Cauchy-Schwartz inequality, the fastest way to decrease energy $KL(\Omega(\tau))$ is obtained by choosing $\frac{\partial C}{\partial \tau} = -F \cdot \mathcal{N}$, where τ is an artificial time and $C = \partial\Omega$, which leads to the evolution equation:

$$\begin{aligned} \frac{\partial C}{\partial \tau} &= \left\{ \int_{\mathbb{R}} \frac{1}{|\Omega|} \left(1 - \frac{q_{out}(I, \Omega)}{q_{in}(I, \Omega)} + \log \frac{q_{in}(I, \Omega)}{q_{out}(I, \Omega)} \right) [K(I - I(s)) - q_{in}(I, \Omega)] dI \right. \\ &\quad \left. + \int_{\mathbb{R}} \frac{1}{|\Omega_0 \setminus \Omega|} \left(1 - \frac{q_{in}(I, \Omega)}{q_{out}(I, \Omega)} + \log \frac{q_{out}(I, \Omega)}{q_{in}(I, \Omega)} \right) [-K(I - I(s)) + q_{out}(I, \Omega)] dI + \lambda \kappa \right\} \mathcal{N}, \end{aligned} \quad (5.3)$$

where the last term $\lambda \kappa$ has been added in the evolution equation in order to regularize the evolving curve. κ is the curvature of the contour C and it is derived from the minimization of the curve length $\int_{\partial\Omega} ds$, ds is the arc length element and λ is a positive constant.

5.2 Existence of Minimizers for the Proposed Model

In this section, we show the existence of minimizer of the variational problem proposed in the previous section by the standard method of calculus of variations.

5.2.1 Our Image Segmentation Model

Our energy Functional is composed of two terms: the first one is the Kullback-Leibler functional, which measures the distance between two pdfs, and the second term is a regularization term based on the curvature.

We assume that the given image I is a function $I \in L^\infty(\Omega_0)$, where Ω_0 is a regular open bounded set of \mathbb{R}^n corresponding to the image domain. We define the set \mathcal{U} of all image regions in Ω_0 , i.e. the set of regular open bounded sets of Ω_0 .

The image segmentation problem proposed in Equation (5.1) Section 5.1 consists in finding a set $\Omega \in \mathcal{U}$ which minimizes the following functional:

$$\begin{aligned} F(\Omega) = & - \underbrace{\int_{\mathbb{R}} \left(q_{in}(I, \Omega) \log \frac{q_{in}(I, \Omega)}{q_{out}(I, \Omega)} + q_{out}(I, \Omega) \log \frac{q_{out}(I, \Omega)}{q_{in}(I, \Omega)} \right)}_{KL(\Omega)} \\ & + \lambda \underbrace{\int_{\partial\Omega} ds}_{L(\Omega)}, \end{aligned} \quad (5.4)$$

where ds is the arc length element, $\partial\Omega$ is the boudary of Ω , $\int_{\partial\Omega} ds$ is the Euclidean length of $\partial\Omega$, $\lambda > 0$ is an arbitrary parameter to control the trade-off between the regularization term and the data-based term. Finally the non-parametric pdfs q_{in} and q_{out} are defined as follows:

$$\begin{cases} q_{in}(I, \Omega) = \frac{1}{|\Omega|} \int_{\Omega} K(I - I(\hat{x})) d\hat{x} \\ q_{out}(I, \Omega) = \frac{1}{|\Omega_0 \setminus \Omega|} \int_{\Omega_0 \setminus \Omega} K(I - I(\hat{x})) d\hat{x} \end{cases},$$

where $\Omega_0 \setminus \Omega$ is the complement set of Ω in Ω_0 .

We look for a set $\Omega \in \mathcal{U}$ which minimizes the region functional (5.4). As it is pointed out in [3, 53], the optimization of the functional (5.4) is difficult to carry out since the set \mathcal{U} of regular domains does not have the structure of a vector space. The variation of the domain is thus done through a family of homeomorphism transformations T (i.e. one-to-one with T and T^{-1} continuous), which allows to differentiate F with respect to Ω and determine a minimization flow in Section 5.1. Thus, one possible approach to prove the existence of minimizers for (5.4) is to express (5.4) in term of the transformation T and look for a minimizer T_* . However, we decide to choose another approach to prove the existence of a minimizer which looks easier. This approach consists, in a first step in re-writing the functional F with the *characteristic function* χ_Ω of the set Ω , defined as follows:

$$\chi_\Omega(x) = \begin{cases} 1 & \text{if } x \in \Omega \in \mathcal{U} \\ 0 & \text{otherwise} \end{cases},$$

and then applying the standard method of the calculus of variations to prove the existence of a minimizer. Functional (5.4) can be expressed w.r.t. χ_Ω :

$$F(\chi_\Omega) = -KL(\chi_\Omega) + \lambda L(\chi_\Omega), \quad (5.5)$$

where

$$L(\chi_\Omega) = \int_{\Omega_0} |\nabla \chi_\Omega| dx, \quad (5.6)$$

and

$$\begin{cases} q_{in}(I, \chi_\Omega) = \frac{\int_{\Omega_0} K(I - I(\hat{x})) \chi_\Omega(\hat{x}) d\hat{x}}{\int_{\Omega_0} \chi_\Omega d\hat{x}}, \\ q_{out}(I, \chi_\Omega) = \frac{\int_{\Omega_0} K(I - I(\hat{x})) (1 - \chi_\Omega(\hat{x})) d\hat{x}}{\int_{\Omega_0} (1 - \chi_\Omega) d\hat{x}}. \end{cases} \quad (5.7)$$

Equation 5.6 introduces the total variation (TV) norm, which will play an important role in the theorem of existence. It is defined according to:

Definition 5.2.1. [36, 42] Let $\Omega \in \mathcal{U}$ and $u \in L^1(\Omega_0)$. The TV norm of the function u is defined as follows:

$$TV(u) = \int_{\Omega_0} |\nabla u| dx = \sup_{\phi \in \Phi} \left\{ \int_{\Omega_0} u \operatorname{div} \phi dx \right\},$$

where $\Phi = \{\phi \in C_0^1(\Omega_0, \mathbb{R}^n) \mid |\phi| \leq 1, \text{ on } \Omega_0\}$ and $C_0^1(\Omega_0, \mathbb{R}^n)$ are the continuously differentiable real functions on Ω_0 .

Moreover,

Definition 5.2.2. [36, 42] A function $u \in L^1(\Omega_0)$ is said to have bounded variation in Ω_0 if its distributional derivative satisfies $TV(u) < \infty$. We define $BV(\Omega_0)$ as the space of all functions in $L^1(\Omega_0)$ with bounded variation. The space $BV(\Omega_0)$ is a Banach space, endowed with the norm:

$$\|u\|_{BV(\Omega_0)} = \|u\|_{L^1(\Omega_0)} + TV(u).$$

We introduce two important theorems that are used in Equation (5.6) and in our theorem of existence.

Theorem 5.2.1. [36, 42] A set $\Omega \in \mathcal{U}$ has finite perimeter if and only if the characteristic function χ_Ω of Ω belongs to $BV(\Omega_0)$. We have

$$Per(\Omega) = TV(\chi_\Omega) = \int_{\Omega_0} |\nabla \chi_\Omega| dx < \infty,$$

and

Theorem 5.2.2. [36, 42] Let $\Omega \in \mathcal{U}$. If $\{u_k\}_{k \geq 1}$ is a bounded sequence in $BV(\Omega_0)$, then there exists a subsequence $\{u_{k_j}\}$ of $\{u_k\}$ and a function $u_\star \in BV(\Omega_0)$, such that $u_{k_j} \rightarrow u_\star$ strongly in $L^p(\Omega_0)$ for any $1 \leq p < n/(n-1)$ and

$$TV(u_\star) \leq \liminf_{k_j \rightarrow \infty} TV(u_{k_j}).$$

We can now state the theorem of the existence of (at least) one minimizer for (5.5):

Theorem 5.2.3. *Our minimization problem*

$$\min_{\chi_\Omega \in BV(\Omega_0)} \left\{ KL(\chi_\Omega) + \lambda L(\chi_\Omega) \right\}, \lambda > 0, \quad (5.8)$$

has a solution in $BV(\Omega_0)$.

Proof. The direct method of the calculus of variations ([2, 25, 99]) is used:

A) Let $\{\chi_{\Omega_k}\}_{k \geq 1}$ be a minimizing sequence of (5.8), i.e.

$$\lim_{k \rightarrow \infty} F(\chi_{\Omega_k}) = \inf_{\chi_\Omega \in BV(\Omega_0)} F(\chi_\Omega).$$

B) Since χ_{Ω_k} is a sequence of characteristic functions of the sets Ω_k , then $\chi_{\Omega_k}(x) \in \{0, 1\}$ - a.e. in Ω_0 . A constant $M > 0$ exists such that $\|\nabla \chi_{\Omega_k}\|_{L^1(\Omega_0)} \leq M$, $\forall k \geq 1$. Therefore, χ_{Ω_k} is a uniformly bounded sequence on $BV(\Omega_0)$. Following Theorem 5.2.2, a subsequence of $\chi_{\Omega_{k_j}}$ that converges to a function χ_{Ω_\star} strongly in $L^1(\Omega_0)$ exists.

C) Taking a minimizing sequence $\chi_{\Omega_{k_j}} \rightarrow \chi_{\Omega_\star}$, it is easy to show that $\lim_{k_j \rightarrow \infty} KL(\chi_{\Omega_{k_j}}) = KL(\chi_{\Omega_\star})$ since $q_{in}(\chi_\Omega)$ and $q_{out}(\chi_\Omega)$ in Equation (5.7) are continuous w.r.t. the $BV(\Omega_0)$ topology. Thus, according to Theorem 5.2.2, we deduce that

$$F(\chi_{\Omega_\star}) \leq \liminf_{k_j \rightarrow \infty} F(\chi_{\Omega_{k_j}}),$$

which implies that

$$F(\chi_{\Omega_\star}) = \min_{\chi_\Omega \in BV(\Omega_0)} F(\chi_\Omega),$$

which means that χ_{Ω_\star} is a minimizer of F among characteristic functions χ_Ω of sets $\Omega \in \mathcal{U}$ of finite perimeter in Ω_0 . It also implies the existence of at least one set Ω_\star , given by $\{x \in \Omega_0 | \chi_{\Omega_\star}(x) = 1\}$, which minimizes (5.4). \square

5.2.2 Generalization of the Existence Theorem 5.2.3

The previous existence theorem can be extended to other variational models such as [49, 53] which have the form:

$$\min_{\Omega} \{F(\Omega) = F_d(q_{in}, q_{out}, \Omega) + \lambda L(\Omega)\}, \quad \lambda > 0 \quad (5.9)$$

where F_d is a data-based functional such as (5.1), where functions q_{in}, q_{out} have to be continuous w.r.t. the BV topology s.a. in Eq. 5.1 and L is the length of Ω . Image segmentation models defined as (5.9), are guaranteed to hold a minimizing solution.

5.3 Dual Formulation and Fast Segmentation

New Convex Functional In the previous section, we have defined an energy functional that we want to minimize as fast as possible. In this section, we propose to redefine the energy functional F to develop a fast numerical minimization scheme, which does not need to re-compute the signed distance function regularly. As it has been said in Section 4.2, this numerical scheme is based on the work of Bresson et al. in [13], who proposed to compute a global minimum to the active contour energy in order to be independent of the initial contour position.

A convex variational model for our segmentation model (Section 5.1) can be developed. Actually it is a particular case of Model 1 presented in Section 4.2. Equation 5.3 can be written in the level set formulation as follows:

$$\frac{\partial \phi}{\partial \tau} = (-\lambda V_{KL} + \nabla \frac{\nabla \phi}{|\nabla \phi|}) |\nabla \phi|, \quad (5.10)$$

where V_{KL} is the speed provided by Kullback-Leibler distance 5.1. Since $|\nabla \phi| > 0$, the steady state of (5.11) is the same as:

$$\frac{\partial \phi}{\partial \tau} = -\lambda V_{KL} + \nabla \frac{\nabla \phi}{|\nabla \phi|}. \quad (5.11)$$

Then ϕ in (5.11) is solution of the variational model:

$$\min_{\phi \in [0,1]} F(\phi) = \int_{\Omega_0} -\lambda V_{KL} \phi + |\nabla \phi|. \quad (5.12)$$

To avoid any confusion with the level set function ϕ , the notation is changed and we are seeking the minimum of the functional $F(u)$ such that:

$$\min_{u \in [0,1]} F(u) = \int_{\Omega_0} -\lambda V_{KL} u + \int_{\Omega_0} |\nabla u|, \quad (5.13)$$

where $\int_{\Omega_0} |\nabla u| =: TV(u)$ is the total variation norm of the function u . Based on Theorem 4.2.1 and supposing $V_{KL} \in C^1(\Omega)$ fixed, the characteristic function $\mathbf{1}_{\Omega_C(\mu)}$, for a.e $\mu \in [0, 1]$ is a global solution of Functional.

We want to emphasize that $V_{KL} := V_{KL}(x, p_{in}(x), p_{out}(x))$ in Theorem 4.2.1 is fixed since the pdfs p_{in}, p_{out} are also fixed. It means that the convexity of the functional F is considered w.r.t. the function u but not w.r.t. p_{in}, p_{out} which will be updated during the minimization process.

Fast Algorithm based on Dual Formulation The variational model (5.13) can be quickly minimized using a dual approach of the TV norm as in [4, 13, 21]. We use a convex regularization of the variational model (5.13) as follows:

$$\min_{v \in [0,1], u} F(u, v) = \int_{\Omega_0} |\nabla u| - \lambda V_{KL} v + \frac{1}{2\theta} (u - v)^2, \quad (5.14)$$

where $\theta > 0$. Since the functional F is convex w.r.t. u, v , its minimizer can be computed by minimizing F w.r.t. u, v separately, and iterating until convergence as in the references

mentioned above. Thus, the following minimization problems are considered:

$$v \text{ being fixed, } \min_u \left\{ TV(u) + \frac{1}{2\theta} \|u - v\|_{L^2}^2 \right\}, \quad (5.15)$$

$$u \text{ being fixed, } \min_{v \in [0,1]} \left\{ \int_{\Omega_0} -\lambda V_{KL} v + \frac{1}{2\theta} (u - v)^2 \right\}, \quad (5.16)$$

- The solution of (5.15) is given by [21]: $u = v - \theta \nabla \mathbf{p}$, is given by

$$\mathbf{p}^{n+1} = \frac{p^n + \delta t \nabla(\operatorname{div} \mathbf{p}^n - v/\theta)}{1 + \delta t |\nabla(\operatorname{div} \mathbf{p}^n - v/\theta)|}, \quad n \in \mathbb{N},$$

- and the solution of (5.16) is given by [13]:

$$v = \min\{\max\{u + \theta \lambda V_{KL}, 0\}, 1\}. \quad (5.17)$$

Details for the numerical scheme are in Section 4.2.

At each iteration p_{in}, p_{out} are updated using the Parzen method given in (5.1).

5.4 Experimental Results

In this chapter, results of our algorithm using different kind of features are presented.

5.4.1 Gray-Value Feature

We start by taking the original image pixel values as feature. Experimental results on synthetic and natural images showed promising performances of our segmentation model. Figure 1 presents four objects having the same intensity mean, the same variance as the background but different higher statistical moments. Figures 5.1(a) and 5.1(c) show the initial and the final active contour, and Figures 5.1(b) and 5.1(d) correspond to the probability densities inside and outside the contour C during the evolution process. Our model has managed to distinguish the four regions from the background, despite equal 1st and 2nd statistical moments, because it is based on probability densities which are different for both regions. As natural image, we have chosen the segmentation of a zebra and a leopard picture which presents textured features. Figures (5.2(a)) and (5.2(c)) show the initial active contour and Figures (5.2(b)) and (5.2(d)) the corresponded result. In the presented results, the segmentation is totally based on the current intensity distribution of the gray scale image partition (object and background).

5.4.2 Intrinsic Geometric Feature

We applied our segmentation algorithm to a set of challenging synthetic and real-world textural images. The first step is the texture features extraction with the descriptor introduced in Chapter 3. The pdfs are estimated on these texture features. The synthetic textural image with the zebra shape, Figure 5.3(a), was generated with the Brodatz data set [15]. The natural textural images, Figures 5.4(d), 5.4(g), 5.4(j), 5.5(a), 5.5(d), 5.5(g), 5.5(j) were taken in the Berkeley segmentation data set [65]. As a comparison with the state-of-the-art techniques, we decided to implement the efficient texture segmentation model

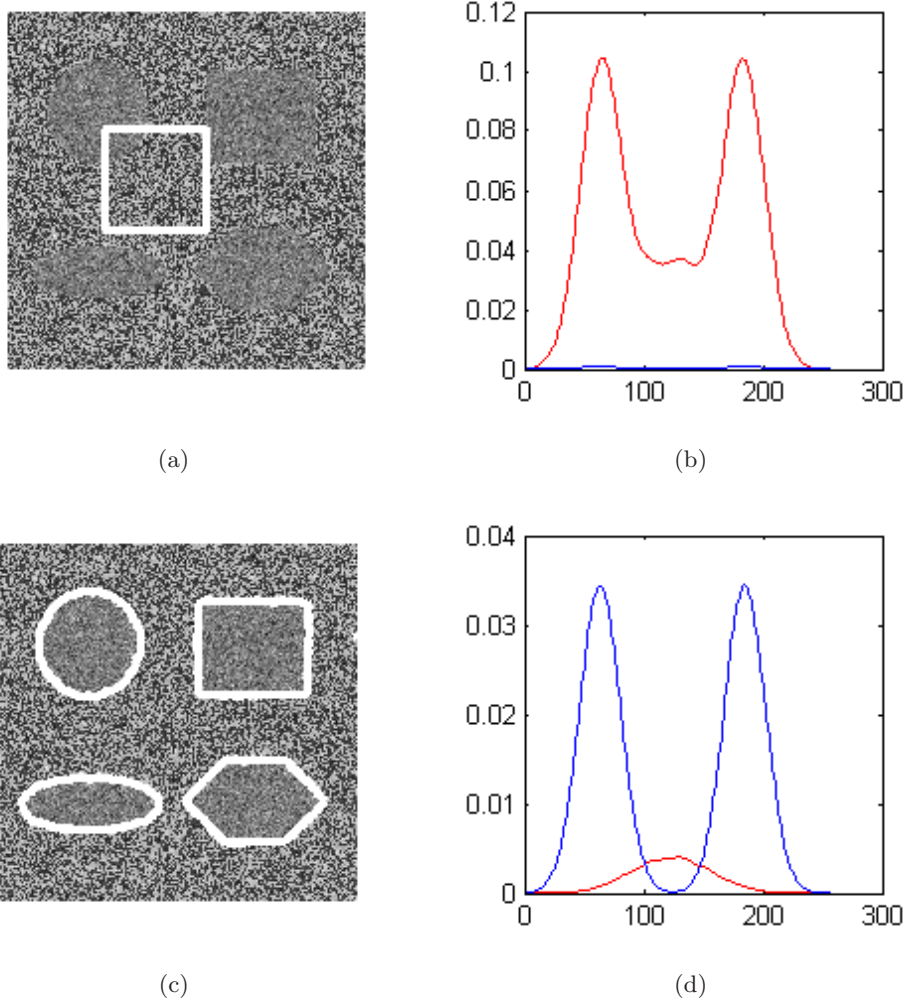


Figure 5.1: Our proposed segmentation model based on KL distance measure. We have considered here the image intensity as a random variable and the KL measure has been defined on the domain of image intensity, i.e. \mathbb{R} . The image represents four objects having the same intensity mean, the same variance as the background but different higher statistical moments. Figures (a-b) present the initialization of the active contour and the associated pdfs inside (blue plot) and outside (red plot) the AC. Figures (c-d) present the final segmentation and the associated pdfs inside (blue plot) and outside (red plot) the AC.

of Savig et al [87], which uses the vectorial Chan-Vese model [23] and an edge detector function based on Gabor responses as explained in Section 2.1. We modified their original model by implementing a dual formulation of their energy functional as done in Section 5.3. Besides, the selected Gabor features are chosen with a simple selection criteria defined in [52] in order to have the most relevant collection of Gabor features. Figure 5.4.2 presents the results obtained with our method on the center column and the model of Sagiv et al. in the right column. We notice that our segmentation model needs three parameters, θ, λ as

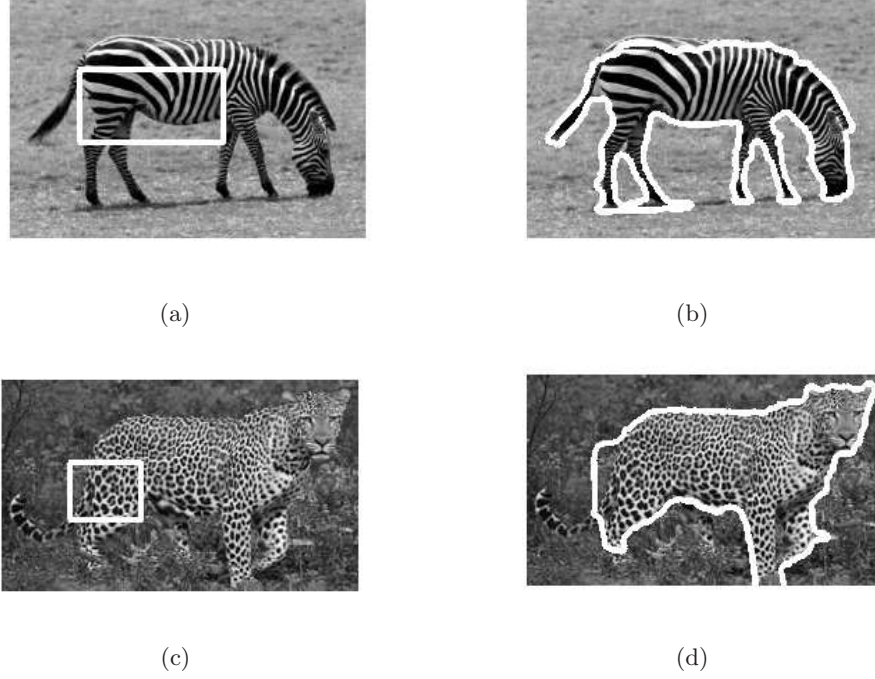


Figure 5.2: Our proposed segmentation model based on KL divergence measure for natural images.

explained in Section 5.3 and σ , the Parzen parameter in Section 5.1. The mean computing time for the segmentation is around a minute. Finally, the method based on Gabor filtering gives comparable results for ordered textures (Figure 5.4.2), but our method significantly improved the results for disordered textures as in natural images. It is particularly clear on the images of the herd of zebras and the sea star, Figures 5.5(g) and 5.5(j).

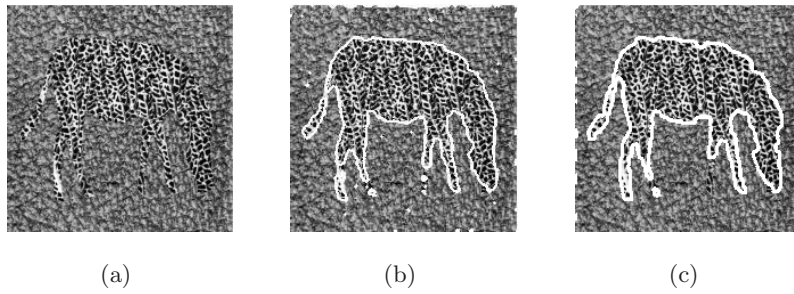


Figure 5.3: Segmentation of synthetic images. Left column: original images. Center column: our segmentation result. Right column: results based on the method [87].

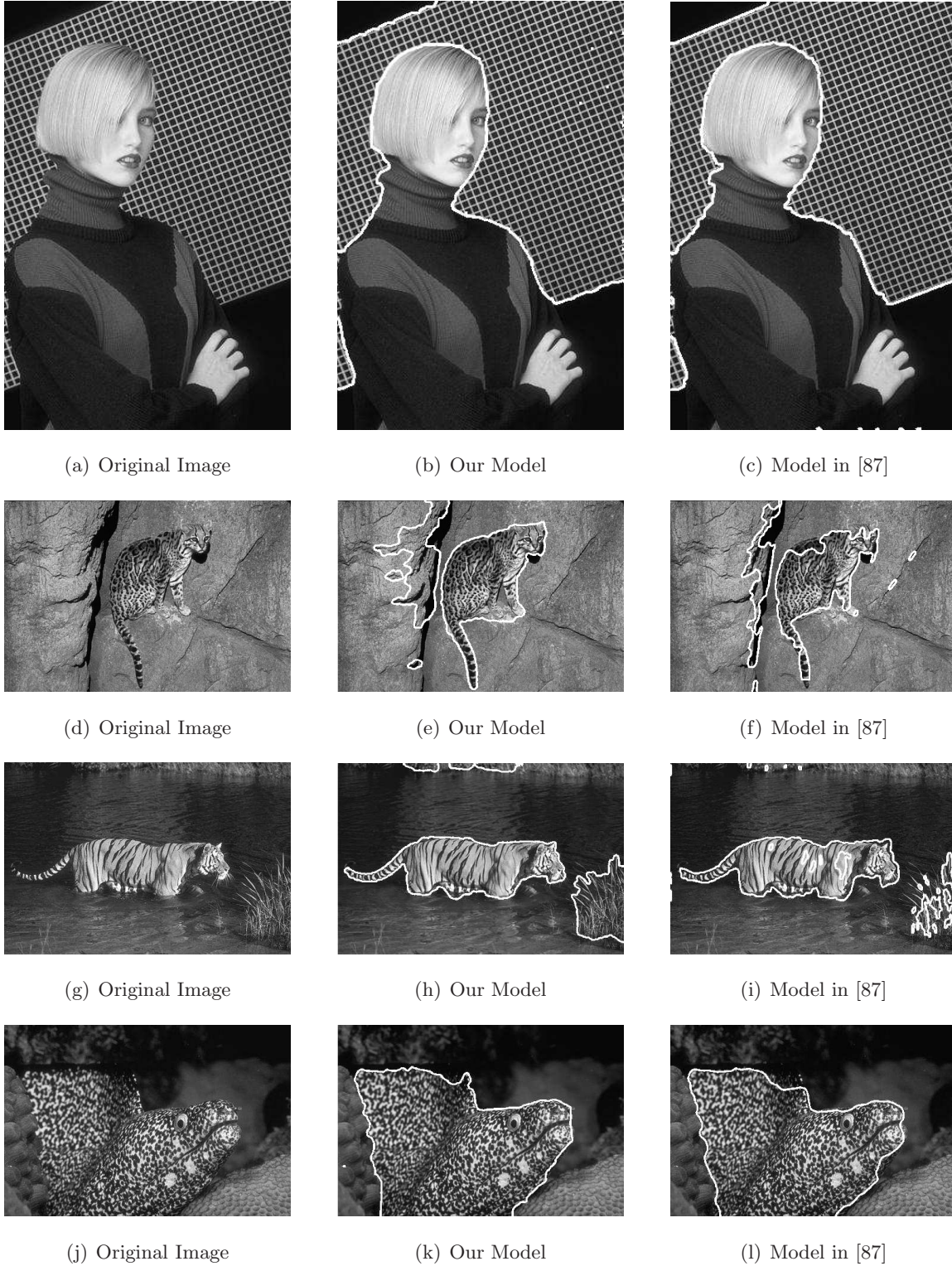


Figure 5.4: Segmentation of real-world textural images. Left column: original images. Center column: our segmentation result. Right column: results based on the method [87].

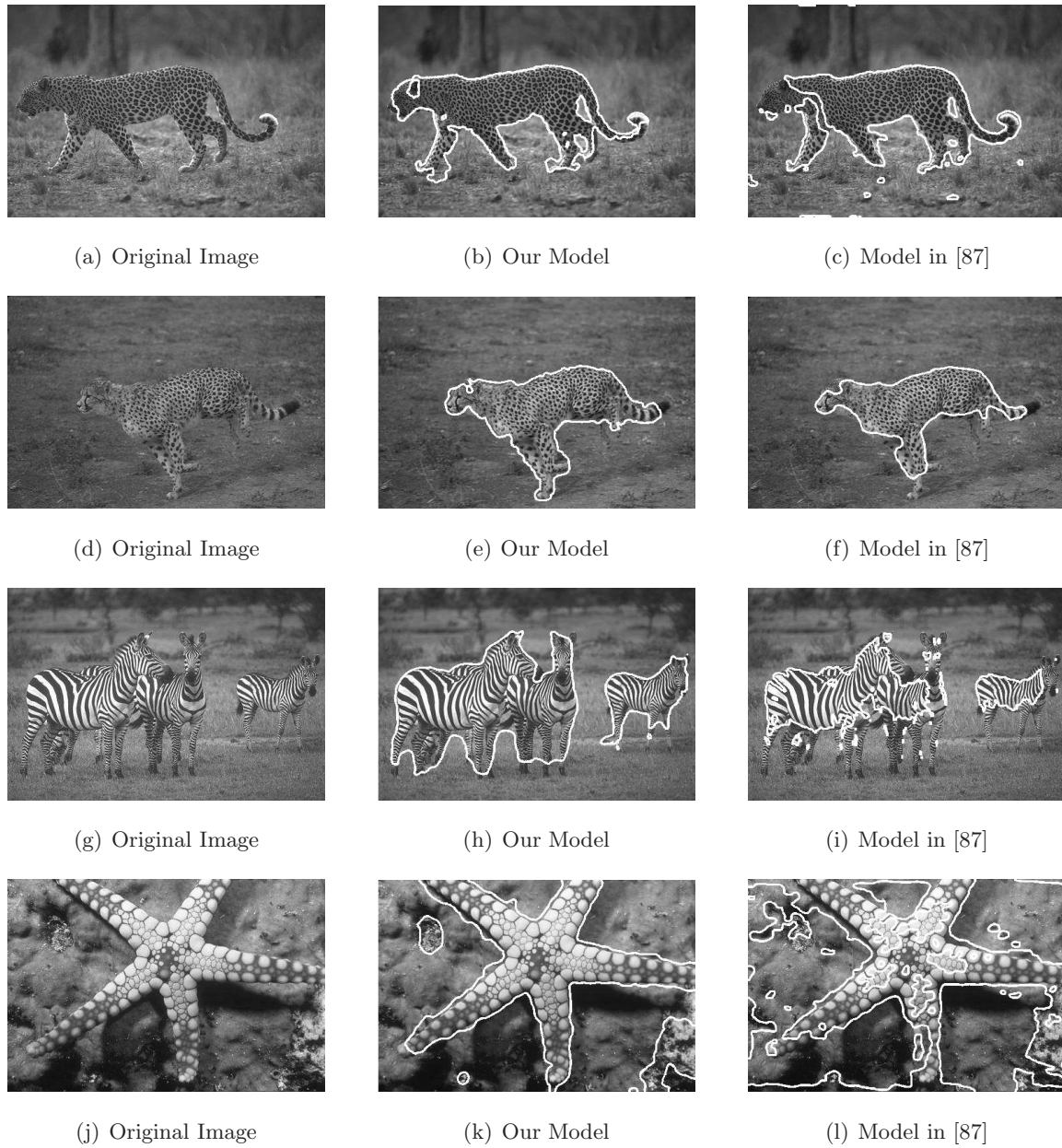


Figure 5.5: Segmentation of real-world textural images. Left column: original images. Center column: our segmentation result. Right column: results based on the method [87].

5.5 Discussion

In this chapter, we have studied the problem of unsupervised image segmentation. Several aspects of the problem have been considered and yield us to use different kind of theoretical tools. From information theory, the Kullback-Leibler (KL) distance has been extracted. Despite of some numerical misbehavior (high computational cost for the pdf logarithm, and case where the pdf is null thus division by zero can occur), this tool has been preferred to many other existent distance measure because of it good theoretical statement and it

invariance under parameter transformation. The Kullback-Leibler divergence has already been used in computer vision field in various way. For instance, Freedman and Zhang [39] used the KL divergence for object tracking. Based on a density probability model, the KL flow allows the matching between the model and the distribution of the current region in the image. Another application is given for Diffusion Tensor-MRI data, by Wang and Vemuri [104] who redefined the KL divergence for a positive definite 2-tensor probability density. The segmentation process is done by minimizing the KL divergence between an inside average tensor and tensors inside the active contour and an outside average tensor and the tensors outside de active contour.

In our case, we have used the KL distance to achieve a region competition between the background and the object. Our segmentation is then unsupervised and based on variational method and active contours. The active contour obtained from the minimization of the energy functional defined in Equation 5.1 is able to capture highly noisy data (Example Figure 5.4.1.

In contrast with other segmentation methods based on comparison of pdfs ([54, 75]), we do not use any prior assumptions on the probability distribution of the object. Indeed, Paragios and Deriche in [75] based their segmentation process on the Gaussian distribution hypothesis of the regions to be segmented. Jehan-Besson et al. [54] use reference histograms of the region of interest and the background to perform the segmentation.

To segment natural or textured images it is nearly unavoidable to have a feature extraction pre-processing. Rousson et al. in [85] used a non-linear diffusion of the structure tensor from which the statistical information was extracted. From the set of feature and the original image, the maximum likelihood was taken as the segmentation criterion. For highly disordered textures or complex natural images this method is quickly limited because of the choice of feature extractor. It is important to note that the choice of the feature extraction is determinant for the potential that can offer a segmentation method. In fact, Gabor filters are appreciate for the quality of the extracted information, only of course if enough of these responses are considered. As it has been said in the first part, our texture descriptor is particularly appropriate for natural images. Our segmentation results for natural images are explained by the quality of the texture descriptor. In fact the proposed segmentation method is dependent on the feature quality. However this drawback is shared by all these types of segmentation approach.

In the next chapter, we will present a multi-phase active contour segmentation model which leads to global solutions.

Global Multi-Phase Segmentation

6

So far we were limited to two phase segmentation problem, i.e. one object and the background. In this chapter, we propose an extension of the global active contour segmentation model [13, 24] to multi-region active contour segmentation. Such generalization is well justified since most of images are composed of several regions.

6.1 Multi-Region Segmentation with Level Set Function

The literature for multi-region segmentation based on level set representation proposes two main approaches. These two approaches lead to a local solution of the segmentation problem as AC methods classically allow. The first one consists of taking the same number of level set function as the expected number of regions in the image, i.e. m level set functions will evolve simultaneously to capture m regions. Each level set is coupled with all the others by an energy term in order to avoid any vacuum (a pixel has not been attributed to none of the regions) or overlapping problem (a pixel attributed to more than one region).

This concept has been introduced by Zhao, Chan, Merriman and Osher [67, 109] for the motion of multiple junctions. Each level set ϕ_i is driven by its curvature flow and constrained to:

$$\sum_{i=1}^N H(\phi_i(x, y)) - 1 = 0, \quad (6.1)$$

where $H(z)$ is the heaviside function defined as

$$H(z) = \begin{cases} 1 & \text{if } z > 0 \\ 0 & \text{otherwise} \end{cases} . \quad (6.2)$$

Each pixel is constrained to be assigned to exactly one region. Overlapping and vacuum problem are thereby avoided. This coupling formulation has been incorporated into the level set segmentation model [76, 88] by mean of the Lagrange multiplier. Samson et

al. [88] supposed N regions on the hole image have a Gaussian distribution. From these assumptions a regularized partition with the same coupled term is searched.

The second approach for level set multi-region segmentation has been proposed by Vese and Chan in [100]. Their approach relies on the fact that from m level set function, 2^m regions can be described. In fact with one level set, two regions can be defined, thus if one considers all the possible overlapping combination between m level set, then 2^m possibilities of regions can be found. Figure 6.1 illustrates this concept for the case where $m = 2$ which gives four possible regions. The number of regions into the image must be a power of 2. On the other hand, the vacuum and overlapping problem is naturally handled. We chose to adapt this formulation of the multiregion AC segmentation in the dual formulation presented in Section 5.3.

Let us now recall the formulation for multiphase segmentation method in [100] for the case $m = 2$. To this end, some notation should be first introduced. A region \mathcal{R}_{ij} is labeled as follows:

$$\begin{aligned}\mathcal{R}_{11} &= \{z \in \Omega \mid \phi_1(z) > 0 \text{ and } \phi_2(z) > 0\} \\ \mathcal{R}_{01} &= \{z \in \Omega \mid \phi_1(z) \leq 0 \text{ and } \phi_2(z) > 0\} \\ \mathcal{R}_{10} &= \{z \in \Omega \mid \phi_1(z) > 0 \text{ and } \phi_2(z) \leq 0\} \\ \mathcal{R}_{00} &= \{z \in \Omega \mid \phi_1(z) \leq 0 \text{ and } \phi_2(z) \leq 0\}\end{aligned}$$

where c_{ij} refers to the intensity mean in region \mathcal{R}_{ij} and H is the heaviside function as defined in Equation 6.2.

The four-phase segmentation in [100] relies on the minimization of the following energy:

$$\begin{aligned}F_{VC}(\mathbf{c}, \Phi) &= \int_{\Omega} (u_0 - c_{11})^2 H(\phi_1) H(\phi_2) dx dy \\ &+ \int_{\Omega} (u_0 - c_{10})^2 H(\phi_1) (1 - H(\phi_2)) dx dy \\ &+ \int_{\Omega} (u_0 - c_{01})^2 (1 - H(\phi_1)) H(\phi_2) dx dy \\ &+ \int_{\Omega} (u_0 - c_{00})^2 (1 - H(\phi_1)) (1 - H(\phi_2)) dx dy \\ &+ \nu \int_{\Omega} |\nabla H(\phi_1)| + \nu \int_{\Omega} |\nabla H(\phi_2)|,\end{aligned}\tag{6.3}$$

where $\mathbf{c} = \{c_{11}, c_{10}, c_{01}, c_{00}\}$ and ν is constant positif weighting parameter. This energy is a an extension of the two-phase case (Section 4.1.2) and seeks at having a partition of the image into four homogenous regions in the sense of mean inside each region. The corresponded Euler-Lagrange equations which drive the active contour represented by the LS functions $\Phi = \{\phi_1, \phi_2\}$, are given as follows:

$$\begin{aligned}\frac{\partial \phi_1}{\partial t} &= \delta(\phi_1) \left\{ \nu \operatorname{div} \left(\frac{\nabla \phi_1}{|\nabla \phi_1|} \right) - [(u_0 - c_{11})^2 - (u_0 - c_{01})^2] H(\phi_2) \right. \\ &\quad \left. + ((u_0 - c_{10})^2 - (u_0 - c_{00})^2) (1 - H(\phi_2)) \right\}, \\ \frac{\partial \phi_2}{\partial t} &= \delta(\phi_2) \left\{ \nu \operatorname{div} \left(\frac{\nabla \phi_2}{|\nabla \phi_2|} \right) - [(u_0 - c_{11})^2 - (u_0 - c_{10})^2] H(\phi_1), \right. \\ &\quad \left. + ((u_0 - c_{10})^2 - (u_0 - c_{00})^2) (1 - H(\phi_1)) \right\}\end{aligned}\tag{6.4}$$

where δ is the dirac function.

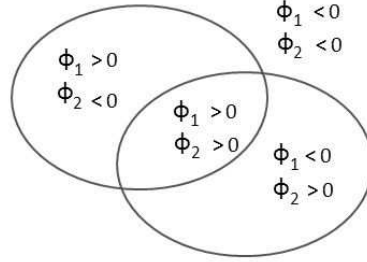


Figure 6.1: The Vese-Chan approach in the case where $m = 2$

6.2 Our New Global Multi-Phase Method

In this section we develop the method to obtain a global minimizer of the multiphase segmentation from the approach of Vese and Chan [100]. We will follow the same steps as in Section 4.2 to construct a segmentation model which solutions are global. A fast segmentation scheme will then solve the problem.

6.2.1 Four Regions Case

Chan and Vese Model: Mean Descriptor. The Euler-Lagrange Equation 6.4 has the same steady state as the following partial differential equation:

$$\frac{\partial \phi_1}{\partial t} = \operatorname{div}\left(\frac{\nabla \phi_1}{|\nabla \phi_1|}\right) - \lambda r_1(x, c_{00}, c_{01}, c_{10}, c_{11})\phi_1 dx, \quad (6.5)$$

$$\frac{\partial \phi_2}{\partial t} = \operatorname{div}\left(\frac{\nabla \phi_2}{|\nabla \phi_2|}\right) - \lambda r_2(x, c_{00}, c_{01}, c_{10}, c_{11})\phi_2 dx, \quad (6.6)$$

where

$$\begin{cases} r_1(\phi_1, \phi_2, \mathbf{c}) = ((I - c_{11})^2 - (I - c_{01})^2)H(\phi_2) + ((I - c_{10})^2 - (I - c_{00})^2)(1 - H(\phi_2)), \\ r_2(\phi_1, \phi_2, \mathbf{c}) = ((I - c_{11})^2 - (I - c_{10})^2)H(\phi_1) + ((I - c_{01})^2 - (I - c_{00})^2)(1 - H(\phi_1)), \end{cases}$$

$\mathbf{c} = (c_{00}, c_{01}, c_{10}, c_{11})$. For sake of simplicity we will write r_i for $r_i(\phi_1, \phi_2, \mathbf{c})$, $i = (1, 2)$. Then Φ is solution of the variational model:

$$\begin{cases} F(\phi_1) = \int_{\Omega_0} -\lambda r_1 \phi_1 + |\nabla \phi_1|, \\ F(\phi_2) = \int_{\Omega_0} -\lambda r_2 \phi_2 + |\nabla \phi_2|. \end{cases} \quad (6.7)$$

As in the previous chapter, we will once again change the symbol of ϕ_i by u_i to avoid any confusion with the level set function and constrain u_1 and u_2 to be in $[0, 1]$, the minimization problem can be written as follows:

$$\left\{ \begin{array}{l} \min_{u_1 \in [0,1]} \left\{ F(u_1) = \int_{\Omega_0} -\lambda r_1 u_1 + |\nabla u_1| \right\} \\ \min_{u_2 \in [0,1]} \left\{ F(u_2) = \int_{\Omega_0} -\lambda r_2 u_2 + |\nabla u_2| \right\} \end{array} \right\}, \quad (6.8)$$

By constraining the u_1 and u_2 to be in $[0, 1]$, the minimization problem becomes convex with respect to u_1 and u_2 for r_1 and r_2 fix.

Region Competition based Probability Density Function. We propose now to use the probability density function as region descriptor. The popular region competition as proposed initially by Zhu and Yuille [110] is used. However we keep a non-parametrical representation of the probability density function as in previous chapters. By keeping the same notation as for the mean descriptor, the region competition between four regions can be described as follows:

$$\begin{aligned}
F(c, \Phi) = & -\left\{ \int_{\Omega} \log q_{11} H(\phi_1) H(\phi_2) dx dy + \int_{\Omega} \log q_{10} H(\phi_1) (1 - H(\phi_2)) dx dy \right. \\
& + \int_{\Omega} \log q_{01} (1 - H(\phi_1)) H(\phi_2) dx dy + \int_{\Omega} \log q_{00} (1 - H(\phi_1)) (1 - H(\phi_2)) dx dy \Big\} \\
& + \nu \int_{\Omega} |\nabla H(\phi_1)| + \nu \int_{\Omega} |\nabla H(\phi_2)|,
\end{aligned} \tag{6.9}$$

where q_{ij} denotes the probability density function in the region \mathcal{R}_{ij} .

$$\begin{aligned}
\frac{\partial \phi_1}{\partial t} &= \delta(\phi_1) \left\{ \nu \operatorname{div} \left(\frac{\nabla \phi_1}{|\nabla \phi_1|} \right) - [(\log q_{11} - \log q_{01}) H(\phi_2) + (\log q_{10} - \log q_{00}) (1 - H(\phi_2))] \right\}, \\
\frac{\partial \phi_2}{\partial t} &= \delta(\phi_2) \left\{ \nu \operatorname{div} \left(\frac{\nabla \phi_2}{|\nabla \phi_2|} \right) - [(\log q_{11} - \log q_{10}) H(\phi_1) + (\log q_{10} - \log q_{00}) (1 - H(\phi_1))] \right\}.
\end{aligned}$$

Then Φ is solution of the variational model:

$$\begin{cases} F(\phi_1) = \int_{\Omega_0} -\lambda s_1 \phi_1 + |\nabla \phi_1|, \\ F(\phi_2) = \int_{\Omega_0} -\lambda s_1 \phi_2 + |\nabla \phi_2|. \end{cases} \tag{6.10}$$

where $s_1(\phi_1, \phi_2) = (\log q_{11} - \log q_{01}) H(\phi_2) + (\log q_{10} - \log q_{00}) (1 - H(\phi_2))$ and $s_2(\phi_1, \phi_2) = (\log q_{11} - \log q_{10}) H(\phi_1) + (\log q_{10} - \log q_{00}) (1 - H(\phi_1))$. By changing the symbol of ϕ_i by u_i and constraining u_1 and u_2 to be in $[0, 1]$, the minimization problem can be written as follows:

$$\begin{cases} \min_{u_1 \in [0,1]} \left\{ F(u_1) = \int_{\Omega_0} -\lambda s_1 u_1 + |\nabla u_1| \right\}, \\ \min_{u_2 \in [0,1]} \left\{ F(u_2) = \int_{\Omega_0} -\lambda s_2 u_2 + |\nabla u_2| \right\}. \end{cases} \tag{6.11}$$

As for the previous paragraph, each minimization problem becomes convex with respect to u_1 and u_2 for s_1 and s_2 fixed.

6.2.2 Generalization for a Global Solution and Fast Algorithm based on Dual Formulation

We propose now to generalize the models for four-regions segmentation to multiple-regions segmentation. We will adopt the same notation as in [100]. In the general case Energy 6.3 can be written as follows:

$$F_{VC_n}(\mu, \Phi) = \sum_{1 \leq \mathcal{R} \leq N} \int_{\Omega} (I - \mu_{\mathcal{R}})^2 \chi_{\mathcal{R}} dx dy + \sum_{1 \leq i \leq m} \nu \int_{\Omega} |\nabla \chi(\phi_i)|. \tag{6.12}$$

where m is the number of level set and $N = 2^m$ is the number of regions, $\mathcal{R} = [1, N]$ is a label for each region, χ is the characteristic function of each label \mathcal{R} and $\mu_{\mathcal{R}}$ is the mean in

the region \mathcal{R} . In the same way we can deduce that the generalization for the multi-phase case for Energy 6.9 can be expressed as follows:

$$F_{pdf_n}(\mathbf{p}, \Phi) = \sum_{1 \leq I \leq N} \int_{\Omega} -\log q_{\mathcal{R}} \chi_{\mathcal{R}} dx dy + \sum_{1 \leq i \leq m} \nu \int_{\Omega} |\nabla \chi(\phi_i)|. \quad (6.13)$$

where q_I is the probability density function of region \mathcal{R} and $\mathbf{q} = [q_1 \dots q_N]$.

In the same way than for the four-phase case we will be able to write a convex functional of $\mathbf{u} = (u_1, \dots, u_m)$:

$$F_i(u_i) = \int_{\Omega} -\lambda \xi_i u_i + |\nabla u_i|, \quad u_i \in [0, 1] \quad (6.14)$$

where ξ is a function of region descriptor (mean or pdf) and of the regions. In the four-regions case, $\xi = r_i$ or $\xi = s_i$. The minimization of Energy 6.14 leads to the same steady-state as Functional 6.13 (resp. 6.12) in the case where ξ is a function of the means (resp. of the pdf). The variational models (6.7, 6.10) and more generally the model 6.14 can be quickly minimized using a dual approach of the TV norm as in [4, 13, 21]. A convex regularization of these variational models can be expressed as follows:

$$\min_{v_i \in [0, 1], u_i} F_i(u_i, v_i) = \int_{\Omega} |\nabla u_i| - \lambda \xi_i v_i + \frac{1}{2\theta} (u_i - v_i)^2, \quad (6.15)$$

where $i = [1, 2^N]$ and $\theta > 0$. Since the functional F_i is convex w.r.t. u_i, v_i , its minimizer can be computed by minimizing F_i w.r.t. u_i, v_i separately, and iterating until convergence as in the references mentioned above. Thus, the following minimization problems are considered:

$$v \text{ being fixed, } \min_{u_i} \left\{ TV(u_i) + \frac{1}{2\theta} \|u_i - v_i\|_{L^2}^2 \right\}, \quad (6.16)$$

$$u \text{ being fixed, } \min_{v_i \in [0, 1]} \left\{ \int_{\Omega} -\lambda \xi_i v_i + \frac{1}{2\theta} (u_i - v_i)^2 \right\}, \quad (6.17)$$

- The solution of (6.16) is given by [21]: $u_i = v_i - \theta \nabla \cdot p$, with:

$$p_i^{n+1} = \frac{p_i^n + \delta t \nabla (\text{div} p_i^n - v_i / \theta)}{1 + \delta t |\nabla (\text{div} p_i^n - v_i / \theta)|}, \quad n \in \mathbb{N}$$

- and the solution of (6.17) is given by [13]:

$$v_i = \min\{\max\{u_i + \theta \lambda \xi_i, 0\}, 1\}. \quad (6.18)$$

The two iteration schemes are straightforward to implement. Implementation details can be found in Section 4.2.

6.2.3 Existence of a Global Solution

Theorem 4.2.1 can be extended to multiple functions:

Theorem 6.2.1. *Suppose that $u_{j, j \neq i}$ fixed, $\xi_i \in C^1(\Omega_0)$ and $\lambda \in \mathbb{R}_+$, if u_{i*} is any minimizer of $F_i(\cdot)$, then for almost every $\mu \in [0, 1]$ we have that the characteristic function*

$$\mathbf{1}_{\Omega_C(\mu) = \{x: u_i(x) > \mu\}}(x),$$

where C is the boundary of the set Ω_C , is a **global minimizer** of $F_i(\cdot, \lambda)$.

Proof. See [13, 24]. \square

We want to emphasize that ξ_i in Theorem 6.2.1 is fixed since the pdfs $q_{\mathcal{R}}$ or the means $c_{\mathcal{R}}$ are also fixed. It means that the convexity of the functional F is considered w.r.t. the function u but not w.r.t. $q_{\mathcal{R}}$ or $c_{\mathcal{R}}$ which will be updated during the minimization process.

6.3 Results

Our method has been tested for both synthetic and real data in the 2D and 3D case.

Mean Measure The mean descriptor is first used on simple images. Our first test image (Figure 6.2(a)) is a noisy synthetic image of size 156×156 . Figure 6.2(b) and 6.2(c) shows the obtained results. Our model can handle noisy images. The result is obtained in less than 2 seconds while for the multi-segmentation with level-sets the results are obtained in 110 seconds. Note that our algorithm is implemented in Matlab and thus could be widely more efficient. Figure 6.2 shows that we can also obtain satisfactory results on a natural image, considering the fact for this kind of image there is no ground truth .

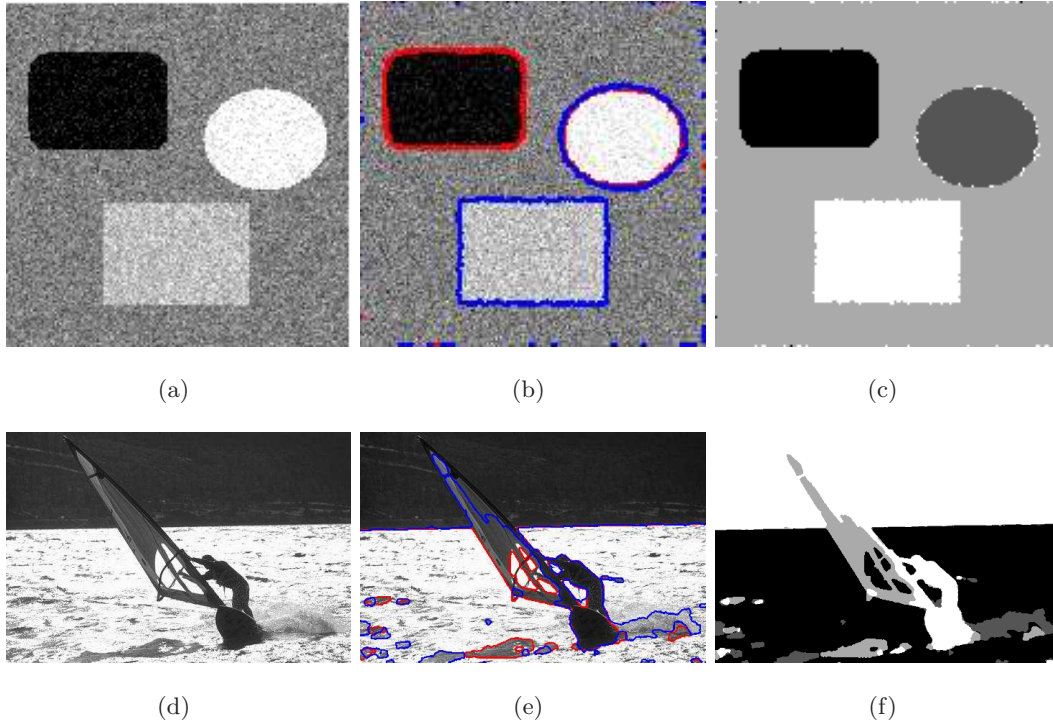


Figure 6.2: Four-phase Segmentation of synthetic and real-world image based on the mean. Left column: original images. Center and right column: our segmentation result.

Pdf Measure Competition of region based on probability density function computed with the Parzen Estimation (Eq. 5.1) is then applied to find four regions in synthetic and real images. Figure 6.3 presents the results that we obtained. We can see that our model can handled synthetic textures as well as natural images from Berkeley data set [65]. In

fact Figure 6.3(a) is composed of four square of synthetic textures and the results of the segmentation is satisfying. For natural images, the task is more difficult since the number of region is difficult to determine. For instance in Figure 6.3(d), luminance variation create an unexpected region. In Figure 6.3(g), the number of objects is higher than four, however the more similar objects have been grouped.

Finally we apply our algorithm to a 3D volume, a Magnetic Resonance Image (MRI) of the brain of size $161 \times 187 \times 161$. The goal is to separate the white matter from the gray matter. (See [28] for more details on the problem). The segmentation result is shown on a axial coronal and sagittal view on Figure 6.4. These are encouraging preliminary results.

6.4 Conclusion

In this chapter a solution for a global minimization of multi-phase active contour segmentation has been proposed. To the best of our knowledge, it is the first time that a global solution for each functional describing the multi-region variational segmentation model has been proposed. We have been using the Chan-Vese model and the region competition between non-parametric pdf. Results on synthetics, real 2D and 3D images have shown the efficiency of our method. Although each variational problem is convex, the joint problem is not convex and this is one of the principal challenge in the future work.

The reason why we did not use the Kullback-Leibler distance as in the previous chapter is for sake of simplicity. Indeed expressing the KL distance in term of the Vese-Chan multi-phase model is not straightforward. Thus this will be a part of the future work.

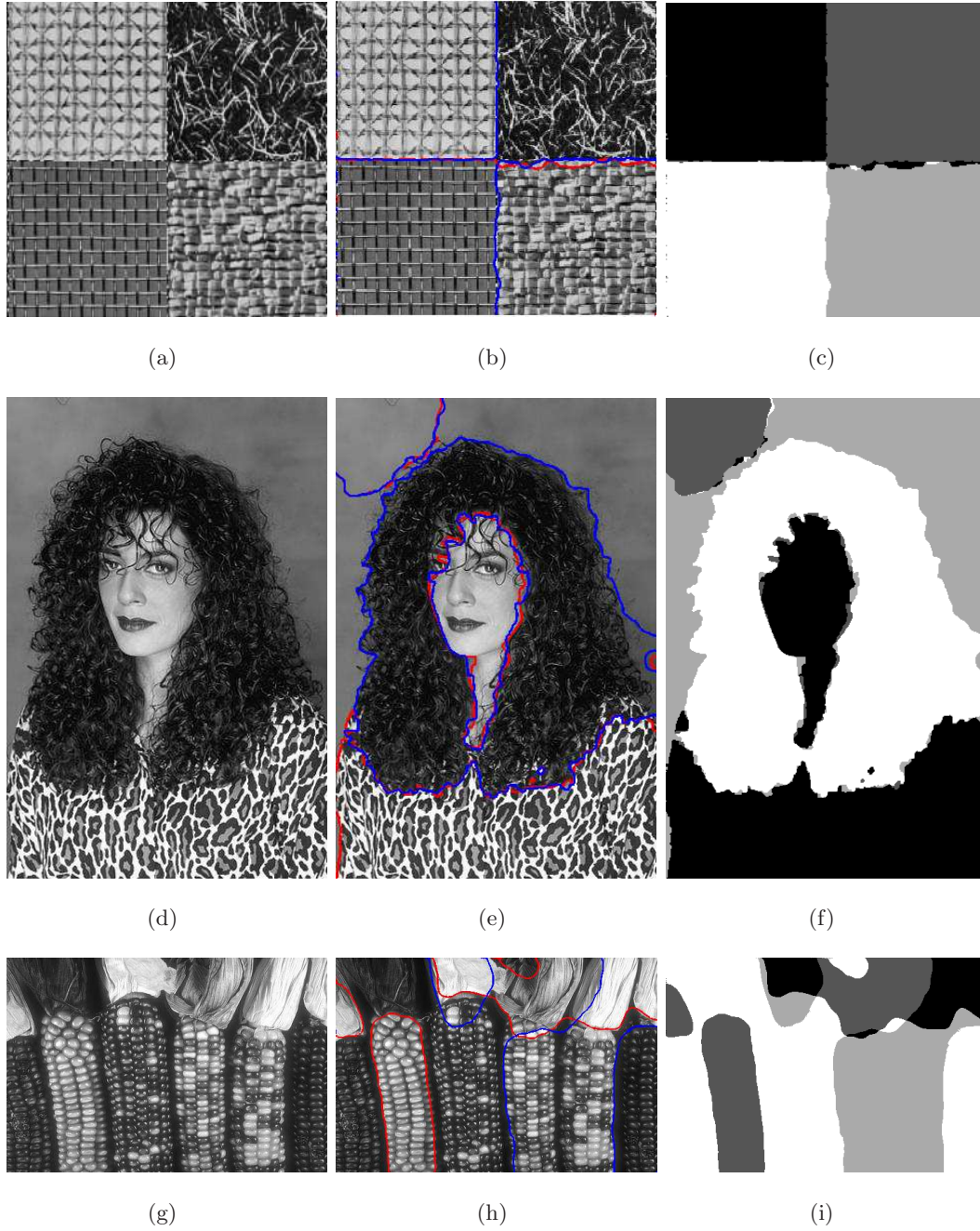


Figure 6.3: Four-phase Segmentation of synthetic and real-world textural images with region competition based on pdf. Left column: original images. Center and right column: our segmentation result.

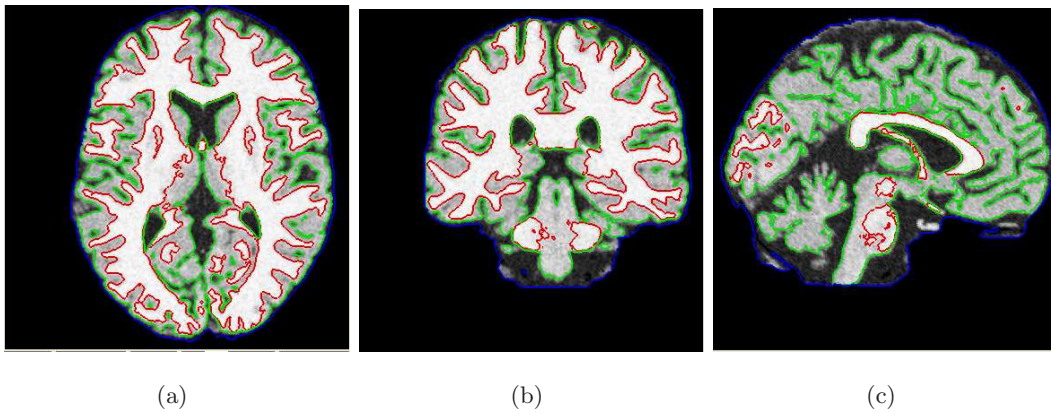


Figure 6.4: Segmentation of the brain (a)Axial view. (b)Coronal view. (c)Sagital view.

Part III

Non-local Semi-Supervised Segmentation Based on Continuous Mincut

Motivations and Contributions

In the third and last part of this thesis, a semi-supervised segmentation algorithm based on a graph non-local image information is presented.

Motivations The proposed segmentation model is based on:

- the use of graph theory: In image segmentation, graph representation is an attractive tool to combine local and non-local image information.
We will see how the segmentation task can be interpreted as a diffusion on graph.
- the use of labels (a hard constraint on the location) to detect objects of interest: In the previous chapter the segmentation process was unsupervised, that means independent from the user. Unsupervised methods offer the comfort of a totally user independent and automatic process. This point of view is attractive however one must face the fact that the image segmentation must deal with different kinds of information, non-information and disinformation. Segmentation methods should handle the huge variety of texture, scale and shape. Moreover most of image segmentation tasks are dependant of what the user calls object of interest (e.g. in medical images). Hence it is sometimes helpful to take into account the user objective.

Contributions In this last part, a new semi-supervised segmentation algorithm has been developed. The contribution is twofold. First, the graph partitioning method is expressed as a variational method. Once we are in this environment, we will show that it is then straightforward to include other variational tools such as for instance the TV-regularization. Second, we introduce hard constraints in the model. The model becomes mathematically well justified. The graph diffusion leads to a well defined segmentation method.

Introduction to Mincut and Semi-supervised Segmentation

7

In this chapter an overview of graph theoretical methods used in image segmentation, in particular for image segmentation, is first given. We will see then how hard constraints on the segmentation results can naturally be incorporated in this framework.

7.1 Graph Diffusion

7.1.1 Image Modeled by a Weighted Graph

In the previous parts of this thesis, an image was defined by a mapping $I : \mathbb{R}^n \rightarrow \mathbb{R}^m$, where n generally equal to 2 or 3 is the image dimension and m is the feature dimension. Variational methods usually suppose that $I \in L^1(\Omega)$. However, in this part we want to extract *non-local information* for each pixel, i.e the relationship one pixel and all the pixels on the whole image. The reason is that non-local information has encountered a large success in various domain of image processing, particularly in image denoising ([18],[40]) and it is particularly well adapted for textures. The non-local information requirement can be achieved by graph representation and the graph theory provides various and powerful methods for processing that kind of data. Historically, graph theory started in 1735 when the Swiss mathematician Leonhard Euler solved a recreational mathematical problem, more precisely the Königsberg bridge problem. There were seven bridges on the Pregel river (Figure 7.1) and the question was : is it possible to walk across all the bridges only once and return to the starting points? The solution consisted of representing each landmass by a node and each bridge as an edge between two landmasses. The graph for this problem was then defined.

A new mathematical area came out and lead to various applications. In fact, many real world situations can be described by graph as instance for social science, chemistry, transportation or communication networks, and then be solved by combinatorial optimization. Graph theory has showed that it could solve several problem in computer vision such as

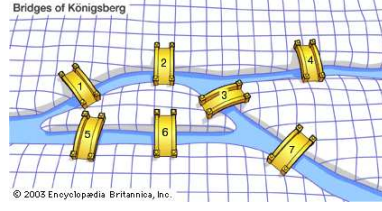


Figure 7.1: Bridges of Konisberg. Figure from [34]

image segmentation [8, 92, 106] and restoration [46], stereovision [102], image and video synthesis [59]. In this domain, the image is modeled by a graph. A graph is a mathematical structure which describes the relationship between elements of a dataset. Each element of this dataset is represented by a node (or vertex) and two nodes are linked by an edge of a certain weight. If the edge between node i and j is equal to the edge between j and i then the graph is said to be undirected, which is the case in this work.

Let $G = (V, E)$ be a weighted undirected graph, where the sets V are the graph nodes and E the edges connection nodes. In the case of image processing, each node V_i represents a pixel/voxel i of an image $I \in \Omega \subset \mathbb{R}^n$. The similarity between two pixels/nodes or nodes i and j is measured by a weight w_{ij} . The graph is then a implicit representation of the image.

The weight matrix W which contains the w_{ij} coefficients has to be carefully defined for the given task.

In image segmentation, two pixels/voxel i and j belonging to the same class should have a high similarity value and in the inverse case the similarity should be null. In practice, we need $w_{ij} = 0$ if node i and node j are "not connected", or in image segmentation terms if pixel i and pixel j do not belong to the same class, and $w_{ij} = 1$ if i and j belong to the same class. A classical construction of weight matrix is based on Gaussian kernel of 0-mean and variance σ . Let $h(i, j)$ be some general distance measure between nodes i and j , then the weight w_{ij} will be computed as follows:

$$w_{ij} = \exp\left(-\frac{h(i, j)}{\sigma^2}\right),$$

where σ is a scaling parameter and $h(i, j)$ is a distance between some local features of the image. The feature can be of any type. Following the definition, the distance function can be:

1. Local, for instance:

$$h(i, j) = \|i - j\|_2. \quad (7.1)$$

The function $h(i, j)$ expresses the spatial proximity of pixel i and pixel j .

2. Non-Local, for instance:

$$h(i, j) = \|F(i) - F(j)\|_2, \quad (7.2)$$

where F is some vector of feature descriptor. The function $h(i, j)$ expresses the proximity between the local descriptors.

3. Combination of Local and Non-Local, for instance:

$$w(i, j) = \exp\left\{-\frac{\|i - j\|_2}{\sigma_1^2} + \frac{\|F(i) - F(j)\|_2}{\sigma_2^2}\right\}. \quad (7.3)$$

In this case the local and non-local function are combined.

The choice of features is once again critical to get an optimal segmentation result. For piecewise smooth images, the gray-level value can be efficient ($F(i) = I(i)$) or for texture images one can for instance consider a feature vector at each pixel coming from a filter bank as suggested in [62] for instance.

An efficient way to represent texture without filter banks is to use the patch around the pixel as feature vector. This idea was introduced first for texture synthesis [32, 61] then, in image denoising, Buades et al. [18] proposed to use patch differences to compute a weight matrix and then perform a non local averaging on a given noisy image. Gilboa et al. [40] proposed a variational model for non-local regularization based on the same weight function. Finally in [12], Bresson et al. proposed a variational unsupervised segmentation method based on this non-local representation of the image.

Let us now introduce a classical and very important operation on the graph, the *minimal cut*.

7.1.2 Mincut or the Diffusion on the Graph

As it has been said previously, the graph is then an implicit representation of the image and several methods allow to perform image processing via the corresponding graph.

For graph-based image segmentation, the problem can be solved by performing a graph partition. By definition a cut is a partition of a graph into two subsets. The general idea of graph partition is to seek for the best subsets $\{U_1..U_k\}$ such that the similarity between U_l and U_k is minimum. We will focus on the two-phase segmentation case, i.e the case where there is one object and one background. The partition is highly dependent on the definition of the weights but also on the definition of a cost function which describes the partition. In graph theory, the optimal partition can be computed by finding the minimal cut (Mincut), i.e. the minimization of the inter-similarity between two sets A and B of V . In other words, given two particular nodes $s \in A$ and $t \in B$ in the graph, the minimum cut is the partition of the graph which gives the minimal capacity defined as the total weight between the nodes for each subset A and B . Most of combinatorial methods for computer vision relies on the Max Flow Min Cut Theorem and a graph representation of the image. The origin of this theorem is network transport. By network transport one should understand any (virtually or really) mobile entity, from oil to human transport passing by network communication. Based on the book of Ford and Fulkerson [37], let us now recall this important theorem. First let $G = (V, E)$ be a network or a graph where the sets V are the graph nodes (points in the network) and E the edges connection nodes also called arcs connection. Two particular nodes are defined in this graph: node s represents the source of the network and t the sink.

A cut C in (V, E) separating s and t is a set of arcs (or edges) (A, B) where $s \in A$, $t \in B$.

Theorem 7.1.1 (Max Flow Min Cut Theorem [37]). *For any network the maximal flow value from s to t is equal to the minimal cut capacity of all cuts separating s and t .*

Algorithmically, the max flow /min cut optimization is handled by combinatorial methods. Description of these methods is out of the scope of this thesis. We just want to specify that these methods are fast and operate on integers numbers which implies a pixel/voxel precision of the results. The interested reader can find more details in [9] and reference therein.

The mincut algorithm has solved different kind of computer vision problem. Greig et al. [46] were the first to use the minimal cut/maximum flow algorithm for image restoration. Wu and Leahy [106] proposed to use the minimal cut algorithm for data clustering. They opened the way to several graph partitioning method. Then Boykov and Jolly [8] adapted the work in [46] in an efficient way for image segmentation. Then Boykov and Kolmogorov [9] proposed a highly efficient and competitive method to resolve the mincut problem. Finally, an other point of view on mincut is taken with the normalized cut method of Shi and Malik [92] where the partitioning problem becomes a spectral problem.

In the segmentation method that will be presented in the next chapter, the starting point is the original mincut definition. We are seeking then for the optimal bi-partition of V into two sets A and B such that $A \cup B = V$ and $A \cap B = \emptyset$. In this case, the minimal cut partition can be written as follows:

$$cut(A, B) = \min_x \sum_{x_i > 0, x_j < 0} -w_{ij}x_i x_j \quad (7.4)$$

where x is a N dimensional indicator vector, with $N = \text{card}(V)$, such that $x_i = 1$ if node $i \in A$, and $x_i = -1$ otherwise.

In the following, we want to outline the direct relationship between partitioning and diffusion. Let W be the symmetric matrix for the weight function where $W(i, j) = w(i, j)$. We can state that the mincut partition (Equation (7.4)) also corresponds to a diffusion process.

Proposition 1. *The graph partition energy defined in Equation 7.4 is equivalent to a Graph Laplacian operator:*

$$cut(A, B) = \frac{1}{4}x^T(D - W)x = \frac{1}{8} \sum w_{ij}(x_i - x_j)^2, \quad (7.5)$$

where D is a $N \times N$ diagonal matrix with $d_i = \sum_j w(i, j)$ on its diagonal, the matrix $D - W$ is called the Graph-Laplacian.

Based on [92]. Let $\frac{x+1}{2}$ (resp. $\frac{x-1}{2}$) be the indicator function for $x_i > 0$ (resp. $x_i < 0$). Equation (7.4) can be written in the matrix form as follows:

$$\begin{aligned} cut(A, B) &= \frac{(1+x)^T}{2} W \frac{(1-x)}{2}, \\ &= \frac{1}{4}(x^T W \mathbf{1} + \mathbf{1}^T W \mathbf{1} - x^T W x - \mathbf{1}^T W x), \\ &= \frac{1}{4}(\mathbf{1}^T W \mathbf{1} - x^T W x), \\ &= \frac{1}{4}x^T(D - W)x. \end{aligned} \quad (7.6)$$

We have used the fact that $x = \{-1, 1\}$ to obtain the last expression.

We have then the following equalities ([103]):

$$\begin{aligned}
 x^T(D - W)x &= \sum_i^N d_i x_i^2 - \sum_{i,j}^N x_i x_j w_{ij} \\
 &= \frac{1}{2} \left(\sum_i^N d_i x_i^2 - 2 \sum_{i,j}^N x_i x_j w_{ij} + \sum_j^N d_j x_j^2 \right) \\
 &= \frac{1}{2} \sum_{i,j} w_{ij} (x_i - x_j)^2
 \end{aligned}$$

Finally we have:

$$cut(A, B) = \frac{1}{8} \sum_{i,j} w_{ij} (x_i - x_j)^2 \quad (7.7)$$

□

The weighted Graph Laplacian corresponds to a finite difference approximation of the continuous lapacian operator. The graph Laplacian is a non-local operator of diffusion on the graph.

The main disadvantage of minimizing cut (7.7) is to favor isolated clusters. This algorithm is thus very sensitive to noise [92]. To overcome this drawback, normalized cuts (NC) [92] have been proposed by Shi and Malik. It is defined as the fraction between the minimal cut and the total edge connection. The problem is then solved with spectral techniques. Besides, since it would not naturally perform smoothing, mincut can handle more easily highly non-convex shapes and, as we will see further, has interesting mathematical properties .

In this work, we will propose the minimal cut bi-partitioning algorithm in a constraint variational framework. It will be then straightforward to handle the noise by adding the the TV-norm which regularizes the geometry of the contour.

Before presenting our method in the next chapter, we will quickly review state of the art in semi-supervised segmentation for both graph-based and continuous based methods.

7.2 Semi-Supervised Segmentation

In the same way than X.Zhu defined semi-supervised learning in [111], semi-supervised segmentation can "refer to the use of both labeled and unlabeled data". The survey on semi-supervised learning in [111] provides an interesting view of some semi-supervised techniques. For semi-supervised or constrained segmentation the data corresponds to the pixels in the image. Labeling some pixels means that hard constraints are put on these pixels to belong to a certain class. Mincut algorithms naturally suggest semi-supervised techniques for the segmentation with the existence of particular nodes, the source "s" and the sink "t". It becomes natural to designate some pixels as source and some other as sink either if the pixel belongs to the object or to the background. The constraints are added by construction of the graph. In graph-based partitioning, several methods ([6, 8, 108]) are based on these hard constraints or bias. Semi-supervised segmentation of 3-D images, including medical images has also been studied and solved as for instance in [7, 44].

An other approach proposed by Szlam et al [96] consisted of diffusing a characteristic function of the labeled classes and to use spectral methods to obtain the classification.

Constrained segmentation process based on a variational method has also been proposed. In [98], Unger, Pock and Bischof incorporated in the active contour framework seed points, provided by the user, assigned either to the object or to the background. Gilboa et al. [40] also proposed a semi-supervised segmentation framework based on initial object and background labels which will follow a continuous diffusion process driven by a non-local weighted equation. The active contour evolved then following a weighted curvature flow and local constraints provided by the labels. An other type of semi-supervised segmentation was proposed by Protiere, Xue and Sapiro [5, 80] with a fast algorithm based uniquely on distance of each unlabeled pixel to the labels. The user provides labels on some part of the object and the background. The unlabeled pixels are assigned to object or the background labels depending on a fast geodesic distance to the labels. The option to correct some misclassification by adding more label is included.

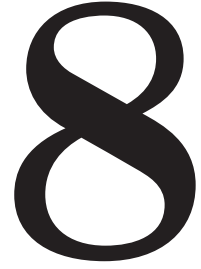
These methods, in particular [40, 98], are directly related to our semi-supervised method that will be presented in the next chapter. In fact we will propose to constraint a segmentation scheme by adding labels assigned to either the object or to the background.

7.3 Conclusion

First of all, this chapter presented the Mincut partitioning method over graphs. We wanted to emphasize the strong link between partitioning and diffusing. Secondly, we noticed how naturally graph partitioning can lead to semi-supervised techniques.

Graph technics are traditionally solved in the discrete space via combinatorial methods. In the next chapter, a semi-supervised segmentation method based on a continuous formulation of the Mincut is presented.

Non-local Semi-Supervised Segmentation based on Continuous Mincut



In this chapter a new non-local semi-supervised method is proposed. The method relies, firstly on a continuous version of the Mincut from graph theory and secondly on the addition of hard constraints in the variational framework.

8.1 Proposed Segmentation Method

8.1.1 Continuous Min Cut (CMC)

We first propose the continuous version of minimal cut problem (7.7), i.e. x_i can take all the values in the interval $[-1, 1]$. Let us make the change of variable $u_i = \frac{x_i+1}{2}$, then $u_i \in [0, 1]$. Since we relax u_i to $[0, 1]$, Eq.(7.7) can be written in the continuous form of as follows:

$$\begin{aligned} E_{CMC}(u) &= \frac{1}{2} \int \int_{\Omega \times \Omega} (u(x) - u(y))^2 w(x, y) dx dy, \quad x, y \in \Omega \times \Omega, \\ &= \frac{1}{2} \int \int_{\Omega \times \Omega} |\nabla_w u|^2 dx dy. \end{aligned} \quad (8.1)$$

where $\nabla_w u$ is the graph gradient of u and $u : \Omega \rightarrow \mathbb{R}$. Minimizing this energy leads us to:

$$\frac{\partial E_{CMC}}{\partial u} = \nabla E_{CMC} = \int_{\Omega} (u(x) - u(y)) w(x, y) dy, \quad (8.2)$$

which corresponds to a weighted Laplacian diffusion on u . The minimization process is done by gradient descent as follows:

$$\frac{\partial u}{\partial t} = - \int_{\Omega} (u(x) - u(y)) w(x, y) dy \quad (8.3)$$

$$= -\nabla E_{CMC} \quad (8.4)$$

Since at each pixel/node x the processing relies on all the other pixels y in Ω , then the method is said to be non-local. We will see that in practice only the pixels in a close neighborhood need to be taken (see Chapter 9).

Functional (8.1) is a particular case of a family of regularizer, the p-dirichlet form, parameterized by the real value p and defined as $\frac{1}{p} \int_{\Omega \times \Omega} |\nabla_w u|^p$. In [33] a study of these regularizers can be found. Let us just precise that we are in the particular case where $p = 2$ and it is also the well-known Tikhonov regularization on weighted graphs.

In [40], G. Gilboa and S. Osher have also studied Functional 8.1 and use it to perform image denoising and semi-supervised segmentation. For performing supervised image segmentation, the regularizing operator is applied on a trinary initial function $u = \{-1, 0, 1\}$. Some marked pixels for the object are assigned to the value 1 and those for the background to the value -1 . The function u is then submitted to the weighted diffusion process via a gradient descent. We propose to improve [40] in two way. Firstly, we define a semi-supervised segmentation model which admits a non trivial solution at final state, while in [40] the solution (u at time $t = \infty$) converges to a constant value $\int_{\Omega} u_0$, where u_0 is the initial condition on u . Secondly, we introduce the TV-norm which regularizes the geometry of the solution.

8.1.2 Semi-Supervised Segmentation with Labels

Our label constrained segmentation will be expressed in a variational framework following the work in [98]. We will show further that these hard constraints lead to a well defined solution of the segmentation problem based on the Mincut Functional (7.4). Suppose that some points of the image are known with some certitude to belong to the object or to the background. For example, a user can select this region manually (with some click). We define a label L_1 as the selected object area and a label L_2 as the selected background area. Let $L(u)$ be the energy describing this constraint:

$$L(u) = \int_{\Omega} \lambda(x)(u - f)^2 dx, \quad (8.5)$$

where $f(x) = \begin{cases} 1 & \text{if } x \in L_1 \\ 0 & \text{if } x \in L_2 \end{cases}$, and $\lambda(x) = \begin{cases} \infty & \text{if } x \in L_1 \text{ or } L_2 \\ 0 & \text{otherwise} \end{cases}$

The function $\lambda(x)$ is the degree of confidence of the data with respect to the labels. Functional (8.5) is defined in order to constrain the function u of being equal to f if $x \in L_1 \cup L_2$ and being equal to anything else outside of $L_1 \cup L_2$. The label constraint energy is obviously highly discontinuous and this discontinuities make the computation of $L(u)$ gradient more complicated. To solve this problem, one can regularize $L(u)$ by considering the functional $L(u, v)$ defined as follows:

$$L(u, v) = \int_{\Omega} \lambda(x)(v - f)^2 dx + \frac{1}{2\theta}(u - v)^2 dx, \quad u : \Omega \rightarrow [0, 1], \quad v : \Omega \rightarrow [0, 1]. \quad (8.6)$$

where θ is chosen small. The gradient of $L(u, v)$ with respect to u and v are given by:

$$\nabla_u L = \frac{1}{\theta}(u - v), \quad (8.7)$$

$$\nabla_v L = 2\lambda(v - f) + \frac{1}{\theta}(u - v). \quad (8.8)$$

The minimization leads to:

$$u = v, \quad (8.9)$$

$$v = \frac{2\lambda(x)\theta f - u}{2\lambda(x)\theta - 1} = \begin{cases} u & \text{if } \lambda = 0 \\ f & \text{if } \lambda = \infty \end{cases} \quad (8.10)$$

8.1.3 Our Semi-Supervised Segmentation Algorithm based on Mincut

The non-local semi-supervised algorithm that we propose is a combination of the weighted diffusion process from Section 8.1.1 and the label constraint from Section 8.1.2. Furthermore, a TV- based regularization on the function u is added. The TV term presents two advantages. Firstly it regularizes the contour between classes. Secondly we explain that Mincut algorithms favor misclassification of small sets, which are smoothed out by a total variation regularization.

The total variation norm is defined by:

$$TV(u) = \int_{\Omega} |\nabla u| dx \quad (8.11)$$

The TV norm has been widely and successfully used in image denoising [86] and image segmentation [13, 24]. The TV norm minimization seeks at minimizing the length of u in the L^1 sense. The TV-norm is highly appreciate in image processing because it disfavors oscillation such as noise while preserving the contrast.

Finally the segmentation energy is as follows:

$$E(u) = \int_{\Omega \times \Omega} |\nabla_w u|^2 dx dy + \int_{\Omega} \lambda(x)(u - f)^2 dx + TV(u). \quad (8.12)$$

and with the regularization of the label energy we get:

$$E(u, v) = \int_{\Omega \times \Omega} |\nabla_w u|^2 + \int_{\Omega} \lambda(x)(v - f)^2 + TV(u) + \frac{1}{2\theta}(u - v)^2. \quad (8.13)$$

In order to use the fast numerical scheme, explained in Chapter 5 Section 5.3, from [13, 21] to minimize $TV(u)$, we add a convex regularization to Functional 8.13:

$$E(u, v, s) = \int_{\Omega} |\nabla_w u|^2 dx dy + \int_{\Omega} \lambda(x)(v - f)^2 dx + TV(s) + \int_{\Omega} \frac{1}{2\theta}(u - v)^2 dx + \int_{\Omega} \frac{1}{2\theta}(s - v)^2 dx. \quad (8.14)$$

The solution of Functional 8.12 is reached by an iterative process done successively over u , v and s as follows:

$$u^{n+1} = u^n - dt(\nabla_u E_{CMC} + \frac{1}{\theta}(u - s)), \quad (8.15)$$

$$v^{n+1} = \begin{cases} \frac{s+u}{2} & \text{if } \lambda = 0 \\ f & \text{if } \lambda = \infty \end{cases} \quad (8.16)$$

$$s^{n+1} = v - \theta \operatorname{div} p^{n+1}, \quad (8.17)$$

where $p = (p^1, p^2)$ is given by

$$p^{n+1} = \frac{p^n + \delta t \nabla(\operatorname{div} p^n) - v/\theta}{1 + \delta t |\nabla(\operatorname{div} p^n) - v/\theta|}, \quad n \in \mathbb{N}$$

See Section (4.2) for more details on the minimization of $TV(u)$.

8.2 Study of the Model

This section deals with the mathematical study of

$$\min_u \{E(u) = \int_{\Omega \times \Omega} |\nabla_w u|^2 dx dy + \int_{\Omega} \lambda(x)(u - f)^2 dx + TV(u)\}. \quad (8.18)$$

A Convex functional The label Energy $L(u)$ and the Mincut Energy are quadratic functions. The total variation of a function $u \in [0, 1]$ has been widely study [24, 86], and it is a convex functional. We can then introduce the following statement:

Proposition 8.2.1. *The Energy $E(u)$ (Eq. 8.18) is a convex functional.*

The convexity of our proposed functional for the segmentation task has important consequences. The convexity implies the existence of a minimizer solution which is more over unique. This is an interesting advantage since the minimization is done via gradient descent and the risk of getting stuck in a local minimum, as it is the case when the functional is non-convex, is discarded. Thus, as far as the labels are correctly defined, the results will be independent of the initialization.

Link with mincut solution Let χ_{Ω_C} be the characteristic function of the set Ω_C defined as follows:

$$\chi_{\Omega_C}(x) = \begin{cases} 1 & \text{if } x \in \Omega_C \subset \Omega \\ 0 & \text{otherwise} \end{cases}$$

Let $u = \chi_{\Omega_C}$, u is then an exact solution of the discrete minimal cut problem (Eq. 7.4):

$$\min_{\Omega_C} \{E_{CMC}(\chi_{\Omega_C}(x))\} = \min_{\Omega_C} \{cut(\Omega_C, \Omega \setminus \Omega_C)\}.$$

In the particular case where u is equal to the characteristic function, the original Mincut problem is recovered.

Existence of a non-trivial solution The image segmentation problem proposed in Section 8.1 consists of finding a continues function u which solves Functional 8.18.

Proposition 8.2.2 (Proposition). *The segmentation formulation defined by Equation 8.18 implies: $u = \overline{u_0}$, where $\overline{u_0} = \int u_0$ and u_0 is the initial condition on u , is not solution of the minimization problem 8.18*

Proof by contradiction. If $u_{\infty} = \overline{u_0}$ then $L(\overline{u_0}) = \int \lambda(x)(\overline{u_0} - f)^2 > 0$.

The minimization implies that $\overline{u_0} = f$.

Except that $f(p) = \begin{cases} 1 & \text{if } p \in L_1 \\ 0 & \text{if } p \in L_2 \end{cases}$

Thus $\overline{u_0} \neq f$.

The solution of 8.18 is given by:

$$u = \begin{cases} \text{anyvalue} & \text{if } \lambda = 0 \\ f & \text{if } \lambda = \infty \end{cases}$$

□

8.3 Conclusion

In this chapter, we proposed our non-local semi-supervised segmentation model. From graph based partitioning method a well defined variational method has been derived. Both, non-local information from the graph and local information from the total variation of the function have been used. All along the process the segmentation is constrained on some part to belong to either the object or the background, which make our segmentation semi-supervised.

The next chapter presents some applications of our method.

Results and Discussion

9

In this chapter, results of our algorithm on different kinds of images are presented. We start with a synthetic image where the influence of the parameters is studied. Then the algorithm is used for object detection in textured and natural images for both gray-value and color images. Finally we apply our algorithm to medical images.

9.1 Experimental Results

The goal here is to study the influence of the different parameters. The graph is constructed from the following local and non-local definition:

$$w(i, j) = \begin{cases} \frac{|i-j|^2}{\sigma_1^2} + \frac{|F(i)-F(j)|^2}{\sigma_2^2} & \text{if } i, j \in \mathcal{N}_{a \times a}(i) \\ 0 & \text{otherwise} \end{cases} \quad (9.1)$$

$\mathcal{N}_{a \times a}(i)$ is square window of size $a \times a$ around i . We suppose thus that if two points are very far there should not be connected. From a^2 neighbors only the cl closest points are kept. The feature vector F is a square patch of size $f \times f$ centered on each pixel. The segmentation is driven by Equations 8.15, 8.16, 8.17. The initial condition for u , v and s are given by the label L_1 , i.e. $u = v = s = 1$ if $x \in L_1$ and $u = v = s = 0$ otherwise.

9.1.1 Test on the Parameters

Our algorithm is applied to a simple image composed of an object and a background with different means disturbed with a Gaussian noise. We propose to study the influence of the different parameters during the graph construction. Our test image is shown on Figure 9.1(a), and the initial label on the object is on Figure 9.1(b). In this simple case we do not need to label the background.

The parameters that will be studied are:

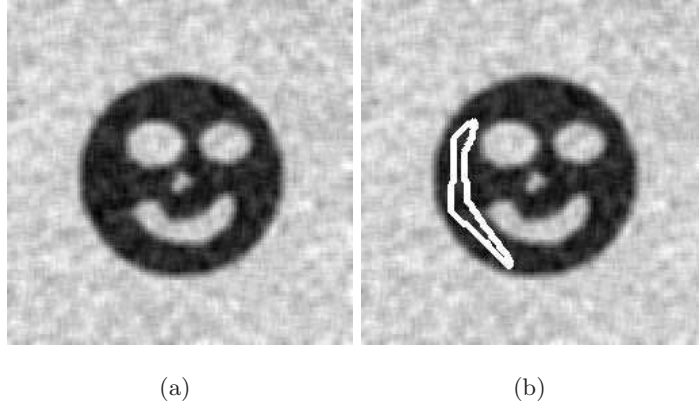


Figure 9.1: (a)Synthetic image test. (b)Initialization with label L_1

- the number of closest neighbors cl ,
- the window search size a ,
- the patch size f ,
- the scale parameter σ_2^2 .

The parameter σ_1^2 has been empirically fixed to be equal to $3f^2$.

Parameter cl We first study the influence of the threshold value cl , for a fixed window search $a = 5$, and a fixed patch size $f = 3$, and a given scale parameter $\sigma_2 = 0.004$. Figures 9.2(a), 9.2(b), 9.2(c) show the result for a number of closest neighbors equal respectively to 2, 3, 24 (which is the total number of neighbors for a window of size 5×5). We can see that for less than 3 neighbors the result is noisy and not satisfactory. However from 3 neighbors to the totality of neighbors the results are quiet similar. The difference will raise in the computational time during the segmentation process. On one hand if the minimal number of neighbors is taken then graph computation will be faster. On the other hand the diffusion process will be slower than if there are a higher number of neighbors. This makes sense because the graph contains the information on the connectivity between points. If each point is already connected to several of its neighbors, then the diffusion process will be faster. We concluded that $cl = 8$ neighbors is a good compromise between time computation for the graph construction and segmentation process. With a unoptimized Matlab implementation, the graph computation lasts approximatively 15 seconds and the segmentation is performed in approximatively 1 minute. The image size is 128×128 .

Parameter a Next the size of the window search is tested. Let us fix $\sigma_2 = 0.004$, $f = 3$. Figure 9.3 shows the segmentation results for two values of a , $a = 5$ and $a = 9$. We notice that details are lost when the window of search get larger.

Parameter f Then the patch size is tested. Figure 9.5 shows the result for a fixed window search of 9 and $\sigma_2 = 0.004$. We can see that when the patch size get bigger the segmentation is coarser and the small details are erased. Of course the size patch depends on the texture

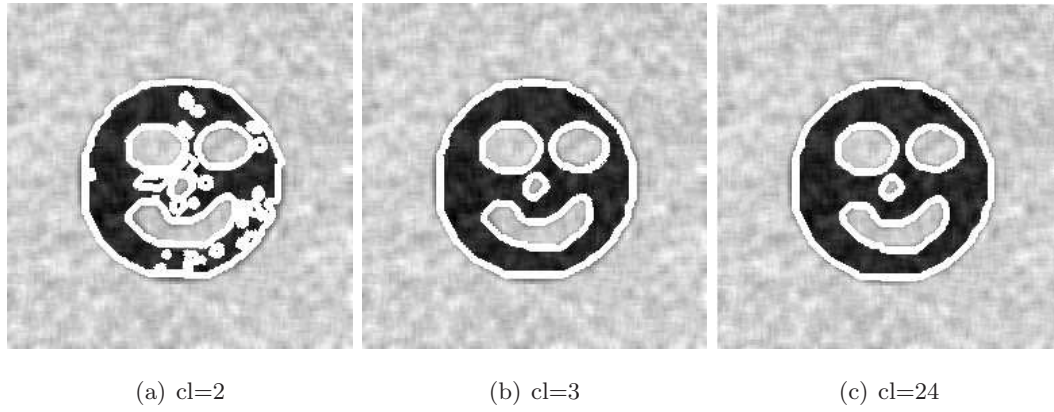


Figure 9.2: Results of our Semi-Supervised Segmentation for different number of selected closest points.

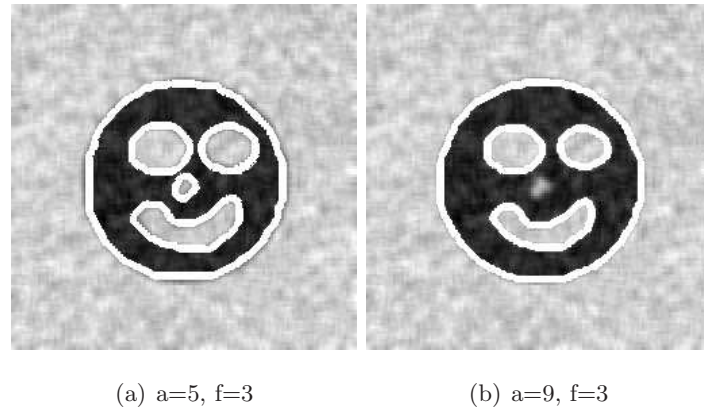


Figure 9.3: Results of our Semi-Supervised Segmentation for different windows size of search

of the object. It should contain a complete texture pattern. This important aspect will be discussed for textures images containing repetitive patterns. If the windows search is too

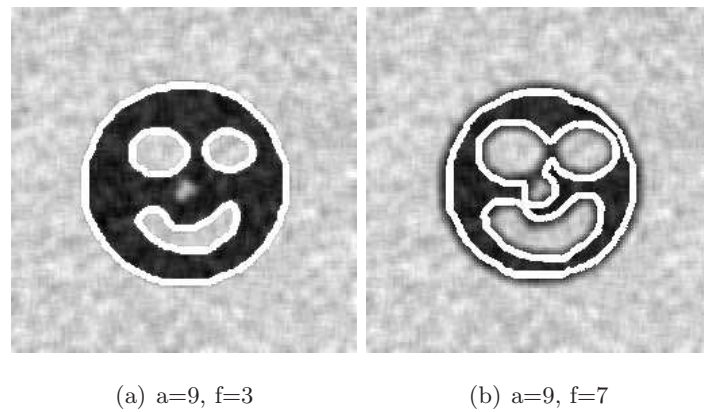


Figure 9.4: Results of our Semi-Supervised Segmentation for different patch sizes

large we notice a loss in details in the segmentation results.

Parameter σ_2 Finally for $a = 5$, $f = 3$, $cl = 8$, the parameter σ_2 is studied. Figure 9.5 shows the different results following the value of σ_2 . A low value of σ_2 implies that only the best pixels are kept. The scale parameter σ_2 should be chosen on the one hand with the goal that enough points are selected to allow the diffusion and on the other hand that the σ_2 value is enough discriminant to select only the pertinent points. Figure 9.5(a) shows the case where σ_2 is too low and Figure 9.5(c) the case where σ_2 is too high.

Finally last very important parameters are the labels. They determine the object to be segmented. Obviously, larger are the labels, easier and faster will be the segmentation.

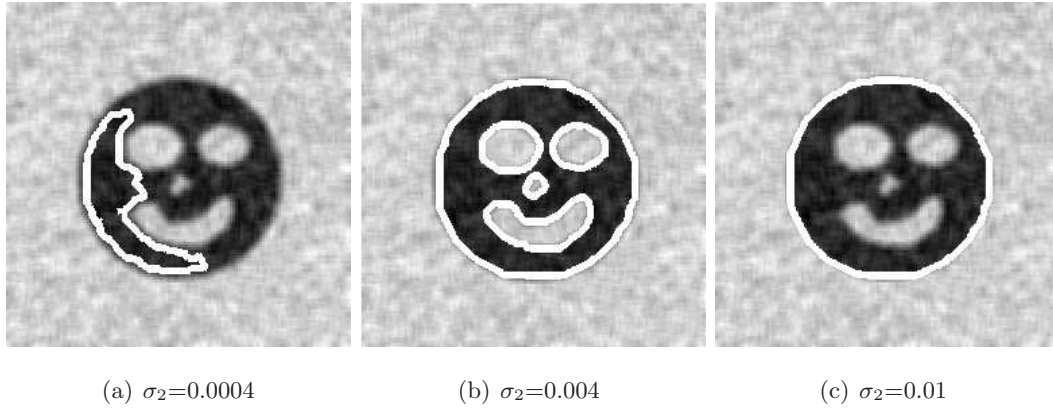


Figure 9.5: Results of our Semi-Supervised Segmentation for three different values of scale parameter.

9.1.2 TV-Regularization Effect

The importance of the TV-Regularization effect is emphasized in this section. A salt and pepper noise is added on a two phase image with different means 9.6(a). The inside and outside labels are shown in Figure 9.6(a). The results shows that if no TV regularization is performed the segmentation fails (Fig 9.6(b)) and when TV regularization is added the segmentation succeeds.

9.1.3 Texture Images

Our algorithm is applied to a synthetic texture image composed of five different patterns 9.7(a). Figures 9.7(b) and 9.7(d) shows the initialization where the wanted texture is chosen and Figures 9.7(c) and 9.7(e) the corresponded results. The size patch is here chosen to be equal to 9×9 which correspond to a pattern size for the two selected textures.

9.1.4 Natural Images

We apply now our algorithm to a set of natural images taken from the Berkeley segmentation dataset [65]. In the first column of Figure 9.8, the inside and outside labels are shown and

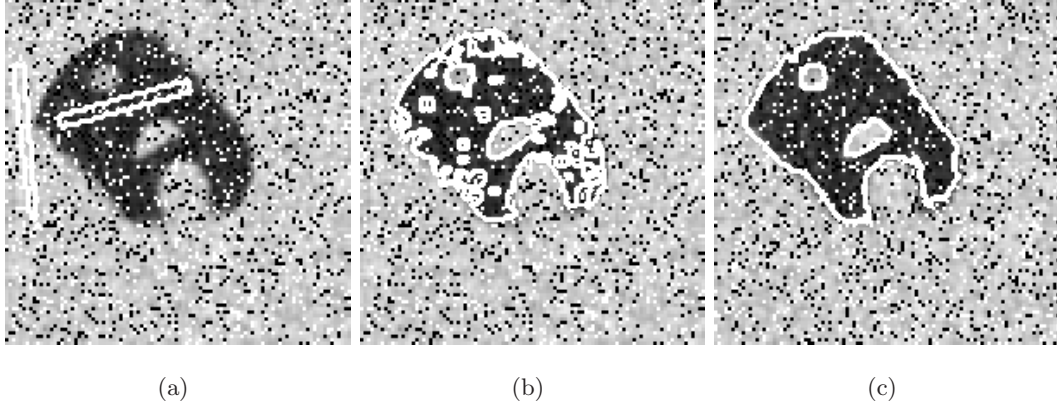


Figure 9.6: Application of our algorithm on a image with a salt and pepper noise. (a)Initialization (b)The segmentation result without a TV-regularization.(c)The segmentation result with a TV-regularization

in the second column the segmentation results.

9.1.5 Color Images

We propose now to apply our segmentation method on color images. Here we consider the simple case of Red-Green-Blue (RGB) channels. The first step consists of computing the graph by taking into account each channel. The distance considered will be expressed as follows:

$$w(i, j) = \begin{cases} \frac{\|i-j\|^2}{\sigma_1^2} + \frac{\|F_r(i)-F_r(j)\|_2 + \|F_g(i)-F_g(j)\|_2 + \|F_b(i)-F_b(j)\|_2}{\sigma_2^2} & \text{if } i, j \in \mathcal{N}_{t \times t}(i) \\ 0 & \text{otherwise} \end{cases} \quad (9.2)$$

where F_r , F_g , F_b are respectively the red, green and blue feature channel. The images are also taken from the Berkeley segmentation dataset [65]. In the first column of Figure 9.8, the inside and outside labels are shown and in the second column the segmentation results.

9.1.6 Medical Images

We apply our segmentation algorithm on 2D medical images of CT scans of the abdomen and the head and neck. Figures 9.10(a) and 9.10(c) shows the inside and outside initial labels. Figures 9.10(b) and 9.10(d) shows the segmentation results. For the liver segmentation, the label on the background (black) prevents the diffusion from capturing as well the heart. The segmentation of the structures in neck are challenging and the result that we obtain is satisfying and encouraging. The extension for 3D images is one of our perspectives for this method.

9.2 Discussion and Conclusion

In this part a non-local semi-supervised segmentation method has been proposed. The success of graph cut in the discrete space for image segmentation was a first motivation

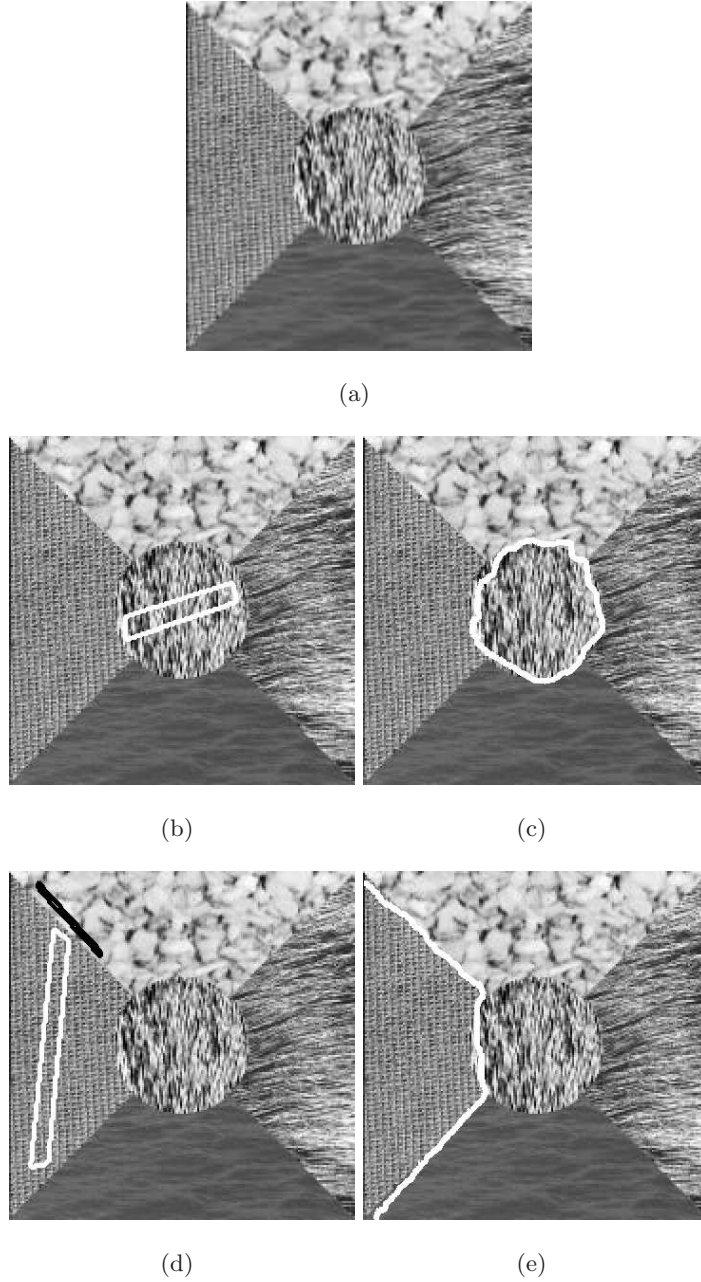


Figure 9.7: Results on synthetic textures. (a)Initial image. (b) and (d) Initialization. (c) and (e) results

to our work. To stay in a variational framework was also set as a goal. By translating the discrete mincut in a continuous way we noticed the direct relationship between diffusion and partition. In fact mincut which is a graph partitioning method has a well know equivalence in the continuous space i.e the Tikhonov regularization. Then how diffusion process which is traditionally a regularizer and denoising method can perform image segmentation? In one sense the graph theory have already answered this question by proving the equivalence between minimal cut and maximal flow of a network. This equivalence give

the intuition of the close relationship between diffusion and segmentation in the continuous space. The continuous formulation of the mincut i.e. a partitioning method is then equivalent to a diffusion process. The addition of hard constraints which made the segmentation semi-supervised allows a solid construction of the segmentation framework.

An other important point for mincut partition is that it relies on graphs which are powerful representations of the non-local information for each pixel. The extraction and representation of features, independently from the segmentation, is a critical issue. Graph representation can give a non-local information, i.e information on a close neighborhood is taken into account.

Our goal was not to outperform combinatorial methods in term of time processing but to take advantage of both the non-local graph information and the tools from variational methods. In work presented local information with TV-norm was used. However, one could use gradient information or even global information with a region competition model. Moreover variational methods provide a sub-pixel/voxel precision whereas combinatorial methods are limited to pixel/voxel precision. This explains faster computational and less precise results of combinatorial methods with respect to variational methods.

In the variational framework, the most related work to ours is probably the method proposed by Gilboa and Osher [40]. The diffusion is driven by the same process however their model was initially drawn for image regularization and can to a trivial solution when adapted to the segmentation task. Our model is well defined for the segmentation and experimental results have shown its efficiency. Future work will focus on extending our method to 3D medical images.

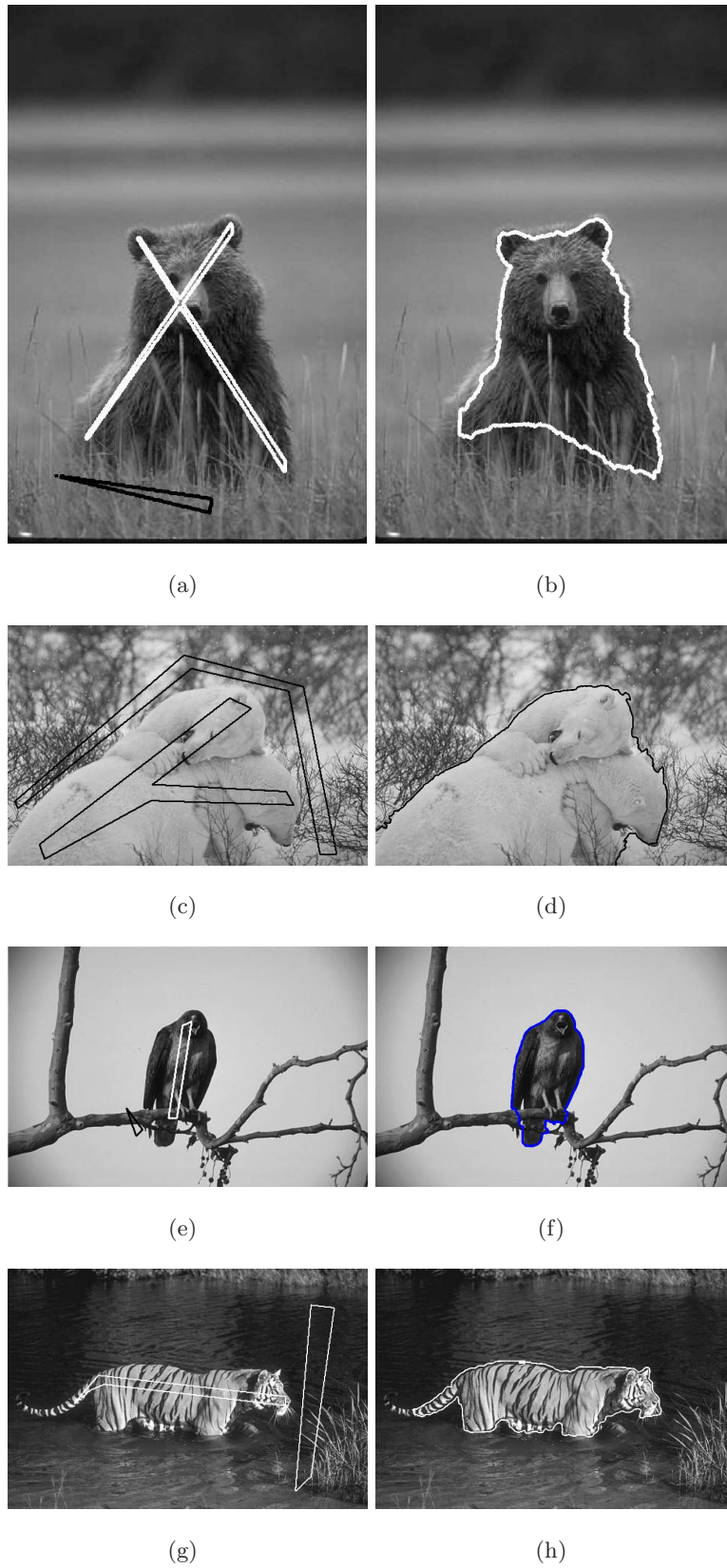


Figure 9.8: Results on real world images from the Berkeley dataset. Left column: Initial labels. Right column: Segmentation Result

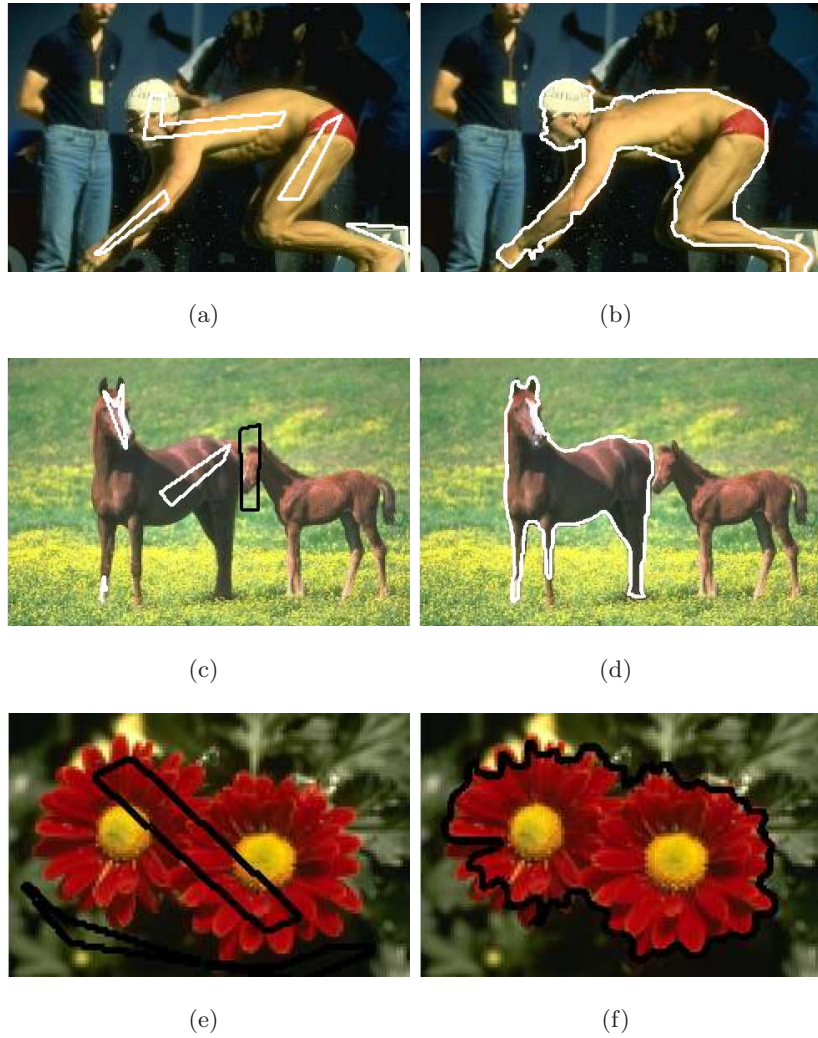


Figure 9.9: Results on real world color images from the Berkeley dataset. Left column: Initial labels. Right column: Segmentation Result

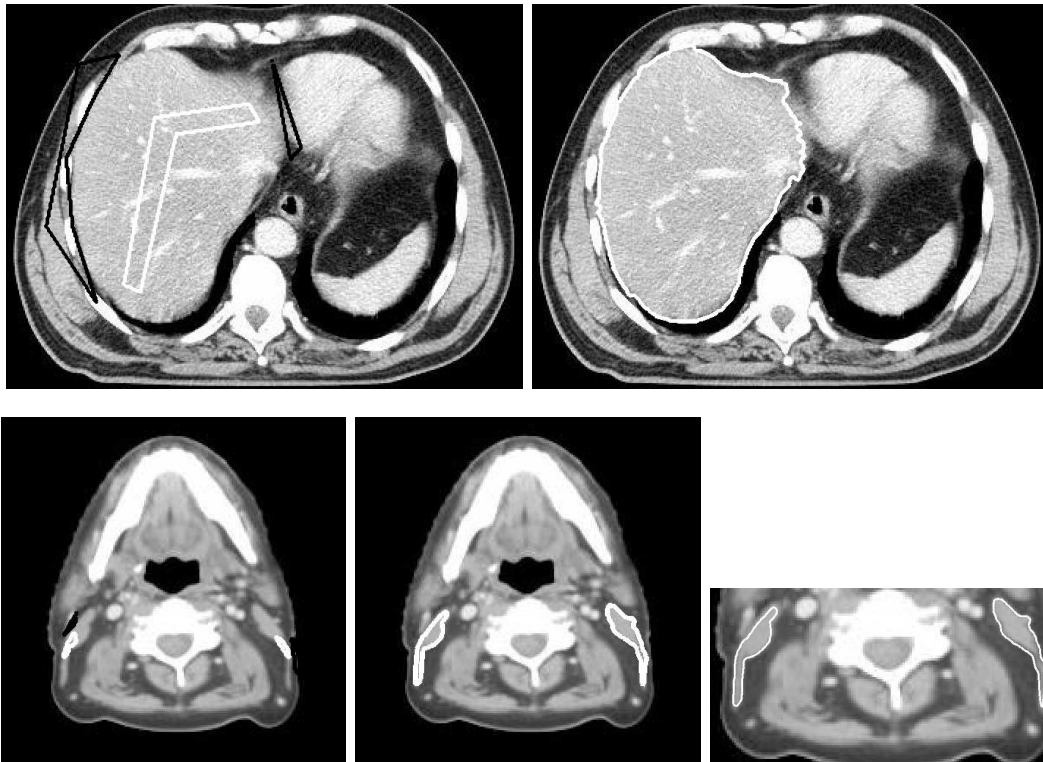


Figure 9.10: First row, Segmentation of the liver. (a)Initial labels. (b)Segmentation Result. Second row, Segmentation of the lateral muscles on the neck. (c)Initial labels. (d)Segmentation Result. (e)Zoom on the segmentation of the muscles

10

Conclusion

10.1 Achievements

The goal of this thesis was to study the problem of object segmentation and to propose algorithms to solve it. Energy based methods solved with variational models were chosen to introduce new efficient segmentation methods for images with textures. The main reason is that these models offer a rigorous mathematical framework, a large flexibility to add any kind of constraints or assumption on the model, including regularization and boundary, region or shape constraints.

The segmentation models that we wanted to introduce meant to be very general, i.e. able to handle any kind of image and particularly textured images. When we tackle the texture image segmentation problem, the extraction of pertinent features arise naturally.

Our first contribution is thus a definition of a new feature for textures. This feature is based on the Beltrami representation of the image and the use of differential geometry. More precisely the image is represented as a surface embedded in \mathbb{R}^3 and the computation of its principal curvatures of this surface are the feature of interest. It has been shown that this feature is particularly efficient for natural images composed of textured regions by one side and smooth or constant regions by the other side. Working on the appropriate feature space makes easier the segmentation task.

Second, An unsupervised region-based active contour segmentation model has been defined. The segmentation is based on the Kullback-Leibler distance and on a non-parametric estimation of the region distribution. Our model is then free of any assumption. A formulation of the problem in a convex functional allowed us to have global minimizers and to use fast numerical schemes to solve the segmentation problem.

This convex formulation has been studied for the multi-region case and our third contribution was to have a convex formulation for the multi-phase active contour segmentation.

This multi-phase segmentation was driven by region competition energy, and a fast implementation scheme was used to solve the problem. In fact compared with the classical level

set segmentation method we noticed that the computational time of our method was 10 times faster. Results on both 2D and 3D images have been shown.

By getting familiar with common applications in image segmentation we get convinced that, in some cases, interaction with the user should be integrated in the framework. The last contribution of this manuscript is a non-local semi-supervised segmentation method. This non-local semi-supervised segmentation stems from graph theory, more precisely from the minimal cut of a graph. A continuous version of the mincut has been incorporate in a variational framework. Tools from variational methods as the TV-norm has been added. Moreover, the graph of image patches provides non-local information on the image and manages, better than local information, the description of regions with textures. In fact textures need to be treated at a non-local level.

10.2 Future Work

A brief description of work in progress and future work is given hereafter:

From the last part of our work, we noticed that the non-local information is powerful to deal with textures. Considering this fact, we would like to examine the idea of a non-local representation of the image with the Beltrami framework conversely to the local representation done in the first part. We could then use tools from differential geometry to extract features of interest in this new manifold.

Concerning the multi-phase segmentation model presented in Part 2, it would be interesting to see the behavior of other multi-phase AC models in the presented convex framework. A formulation based on the Kullback-Leiber distance could be incorporated in the multi-phase segmentation process.

An important issue that have not been studied during this thesis is the multi-scale approach. For improving the time cost of the different segmentation task, in particular for 3D medical volumes, a coarse-to-fine segmentation should be considered.

Finally one important improvement will be in the algorithmic point of view for both unsupervised and semi-supervised segmentation model. We want to consider a recent algorithm for having very fast minimizers of the TV-norm, the Split Bregman algorithm [43].

During this thesis, a large variety of area has been investigated in order to be incorporated in a segmentation framework. Differential geometry, information theory, graph methods have just as much provided us means for constructing segmentation methods. There are still so many others areas to explore for solving this relatively new problem.

Annex



A.1 Shape Derivation Tool

We remind hereafter the shape derivative tool proposed by Delfour and Zolesio in [30]. The shape derivative tool basically derives a region-based functional in an elegant and straightforward way. Let us consider a general region-based functional which depends on an artificial time τ as follows:

$$F(\Omega(\tau)) = \int_{\Omega} f(x, \Omega(\tau)) dx, \quad (\text{A.1})$$

that we want to optimize w.r.t Ω . The Gâteaux derivative of the functional F in the directions V is given by:

$$\langle F'(\Omega), V \rangle = \lim_{\tau \rightarrow 0} \frac{F(\Omega(\tau)) - F(\Omega)}{\tau}, \quad (\text{A.2})$$

where $F'(\Omega) := \frac{\partial F}{\partial \tau}(\Omega(\tau))$. As it has been shown in [3], the Gâteaux derivative can be expressed as:

$$\langle F', V \rangle = \int_{\Omega} \frac{\partial f}{\partial \tau}(x, \Omega, V) dx - \int_{\partial \Omega} f(s, \Omega)(V(s) \cdot \mathcal{N}(s)) ds \quad (\text{A.3})$$

where \mathcal{N} is the unit inward normal to $\partial \Omega$ the boundary of the evolving region Ω , ds its length/area element. In the case where we have $f(x) = f(x, G_1(\tau), G_2(\tau))$ and

$$\begin{cases} G_1(\tau) &= \frac{H_1(\tau)}{H_2(\tau)}, \\ G_2(\tau) &= \frac{H_3(\tau)}{H_4(\tau)}, \end{cases}$$

we obtain:

$$\begin{aligned}\frac{\partial f}{\partial \tau} &= \frac{\partial f}{\partial G_1} \frac{\partial G_1}{\partial \tau} + \frac{\partial h}{\partial G_2} \frac{\partial G_2}{\partial \tau} \\ &= \frac{\partial f}{\partial G_1} \left[\frac{\partial G_1}{\partial H_1} \frac{\partial H_1}{\partial \tau} + \frac{\partial G_1}{\partial H_2} \frac{\partial H_2}{\partial \tau} \right] \\ &+ \frac{\partial f}{\partial G_2} \left[\frac{\partial G_2}{\partial H_3} \frac{\partial H_3}{\partial \tau} + \frac{\partial G_2}{\partial H_4} \frac{\partial H_4}{\partial \tau} \right].\end{aligned}$$

Gâteaux Derivative of Functional KL

$$\begin{aligned}KL(q_{in}(\Omega), q_{out}(\Omega)) &= \int_{\mathbb{R}} (q_{in}(I, \Omega) - q_{out}(I, \Omega)) \\ &\quad \left(\log q_{out}(I, \Omega) - \log q_{in}(I, \Omega) \right) dI.\end{aligned}\tag{A.4}$$

$$\begin{aligned}q_{in}(I, \Omega(\tau)) &= G_1(\tau) = \frac{\int_{\Omega(\tau)} K(I(x) - I(\hat{x})) d\hat{x}}{\int_{\Omega(\tau)} d\hat{x}} = \frac{H_1(\tau)}{H_2(\tau)}, \\ q_{out}(I, \Omega(\tau)) &= G_2(\tau) = \frac{\int_{\Omega_0 \setminus \Omega(\tau)} K(I(x) - I(\hat{x})) d\hat{x}}{\int_{\Omega_0 \setminus \Omega(\tau)} d\hat{x}} = \frac{H_3(\tau)}{H_4(\tau)},\end{aligned}$$

$$\text{and } h(\tau) = G_1(\tau) \log\left(\frac{G_1(\tau)}{G_2(\tau)}\right) + G_2(\tau) \log\left(\frac{G_2(\tau)}{G_1(\tau)}\right),$$

$$\begin{aligned}\frac{\partial h}{\partial \tau} &= \frac{\partial h}{\partial G_1} \left[\frac{1}{H_2} \left(- \int_{\partial\Omega(\tau)} K(I(x) - I(s))(V.N) ds \right) \right. \\ &+ \left. \left(- \frac{H_1}{H_2^2} \right) \left(- \int_{\partial\Omega(\tau)} (V.N) ds \right) \right] \\ &+ \frac{\partial h}{\partial G_2} \left[\frac{1}{H_4} \left(\int_{\partial\Omega(\tau)} K(I(x) - I(s))(V.N) ds \right) \right. \\ &+ \left. \left(- \frac{H_3}{H_4^2} \right) \left(\int_{\partial\Omega(\tau)} (V.N) ds \right) \right].\end{aligned}$$

We can deduce that

$$\begin{aligned}\frac{\partial h}{\partial \tau} &= \frac{1}{|\Omega|} \left(1 - \frac{q_{out}(I(x), \Omega)}{q_{in}(I(x), \Omega)} + \log \frac{q_{in}(I(x), \Omega)}{q_{out}(I(x), \Omega)} \right) \\ &\quad \left(\int_{\partial\Omega(\tau)} (-K(I(x) - I(s)) + q_{in}(I(x), \Omega))(V.N) ds \right) \\ &+ \frac{1}{|\Omega_0 \setminus \Omega|} \left(1 - \frac{q_{in}(I(x), \Omega)}{q_{out}(I(x), \Omega)} + \log \frac{q_{out}(I(x), \Omega)}{q_{in}(I(x), \Omega)} \right) \\ &\quad \left(\int_{\partial\Omega(\tau)} (K(I(x) - I(s)) - q_{out}(I(x), \Omega))(V.N) ds \right)\end{aligned}$$

Finally, the Gâteaux derivative of Functional A.4 is:

$$< KL', V > = \int_{\mathbb{R}} \frac{\partial h}{\partial \tau}(x, \Omega, V) dx,$$

Bibliography

- [1] G. Aubert and L. Blanc-Féraud. Some remarks on the equivalence between 2d and 3d classical snakes and geodesic active contours. *International Journal of Computer Vision*, 34(1):19–28, 1999.
- [2] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. Springer, 2001.
- [3] G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson. Image Segmentation using Active Contours: Calculus of Variations or Shape Gradients. *SIAM Applied Mathematics*, 63(6):2128–2154, 2003.
- [4] J.F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-Texture Image Decomposition - Modeling, Algorithms, and Parameter Selection. *International Journal of Computer Vision*, 67(1):111–136, 2006.
- [5] X. Bai and G. Sapiro. A geodesic framework for fast interactive image and video segmentation and matting. In *IEEE 11th International Conference on Computer Vision*, pages 1–8, 2007.
- [6] A. Blum and S. Chawla. Learning from labeled and unlabeled data using graph min-cuts. In *Proceedings of the Eighteenth International Conference on Machine Learning*, pages 19–26, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.
- [7] Y. Boykov and G. Funka-Lea. Graph cuts and efficient n-d image segmentation. *Int. J. Comput. Vision*, 70(2):109–131, 2006.
- [8] Y. Boykov and M. P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images. volume 1, pages 105–112, 2001.
- [9] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, 2004.
- [10] Y. Boykov and V. Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *Proceedings of the Ninth IEEE International Conference on Computer Vision*, Washington, DC, USA, 2003. IEEE Computer Society.
- [11] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23: 2001, 2001.

-
- [12] X. Bresson and T.F. Chan. Non-local Unsupervised Variational Image Segmentation Models, UCLA CAM Report 08-67, 2008.
 - [13] X. Bresson, S. Esedoglu, P. Vanderghelynst, J.P. Thiran, and S. Osher. Fast global minimization of the active contour/snake model. *Journal of Mathematical Imaging and Vision*, 28(2):151–167, June 2007.
 - [14] C.R. Brice and C.L. Fennema. Scene Analysis Using Regions. *Artificial Intelligence*, 1(3):205–226, 1970.
 - [15] P. Brodatz. Textures: A photographic album for Artists and Designers. New York, NY, Dover Publication, 1996.
 - [16] T. Brox and J. Weickert. A TV flow based local scale estimate and its application to texture discrimination. *Journal of Visual Communication and Image Representation*, 17(5):1053–1073, October 2006.
 - [17] T. Brox, J. Weickert, B. Burgeth, and P. Mrázek. Nonlinear structure tensors. *Image Vision Comput.*, 24(1):41–55, 2006.
 - [18] A. Buades, B. Coll, and J.M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, 4(2):490–530, 2005.
 - [19] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.
 - [20] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic Active Contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.
 - [21] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1-2):89–97, 2004.
 - [22] T.F. Chan and L.A. Vese. Active Contours Without Edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.
 - [23] T.F. Chan, B.Y. Sandberg, and L.A. Vese. Active contours without edges for vector-valued images. *J. Visual Communication Image Representation*, 11(2):130–141, June 2000.
 - [24] T.F. Chan, S. Esedoglu, and M. Nikolova. Algorithms for Finding Global Minimizers of Image Segmentation and Denoising Models, UCLA CAM Report 04-54, 2004.
 - [25] Y. Chen, H.D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K.S. Gopinath, R.W. Briggsand, and E.A. Geiser. Using Prior Shapes in Geometric Active Contours in a Variational Framework. *International Journal of Computer Vision*, 50(3):315–328, 2002.
 - [26] L. D. Cohen, E. Bardinet, and N. Ayache. Surface reconstruction using active contour models. Research report 1824, INRIA, February 1993.
 - [27] D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. *Int. J. Comput. Vision*, 72(2):195–215, 2007.

-
- [28] M. Bach Cuadra, L. Cammoun, T. Butz, O. Cuisenaire, and J.P. Thiran. Comparison and validation of tissue modelization and statistical classification methods in t1-weighted mr brain images. *IEEE Transactions on Medical Imaging*, 25(2):241, 2006.
 - [29] G. Damiand and P. Resch. Split-and-merge algorithms defined on topological maps for 3d image segmentation. *Graph. Models*, 65(1-3):149–167, 2003.
 - [30] M.C. Delfour and J.P. Zolésio. *Shapes and Geometries: Analysis, Differential Calculus, and Optimization*. Advances in Design and Control, SIAM, 2001.
 - [31] L. Devroye. *A course in density estimation*. Birkhauser Boston Inc., Cambridge, MA, USA, 1987.
 - [32] A.A. Efros and W.T. Freeman. Image quilting for texture synthesis and transfer. In *Proceedings of the Conference on Computer graphics and interactive techniques, SIGGRAPH*, pages 341–346, New York, NY, USA, 2001. ACM.
 - [33] A. Elmoataz, O. Lezoray, and S. Bougleux. Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing. *IEEE Transactions on Image Processing*, 17:1047–1060, July 2008.
 - [34] Britannica Online Encyclopedia. <http://www.britannica.com/EBchecked/topic/242012/graph-theory>, 2008.
 - [35] C.L. Epstein and M. Gage. *The curve shortening flow, Fluid Mechanics, Computer Vision and Material Sciences*. Springer-Verlag, New York, 1987.
 - [36] L.C. Evans and R.F. Gariepy. *Measure Theory and Fine Properties of Functions*. Studies in advanced Mathematics, CRC Press, 1992.
 - [37] L.R. Ford and D.R. Fulkerson. *Flows in Networks*. Princeton University Press, Princeton, NJ, 1962.
 - [38] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *ISPRS Intercommission Conference on fast processing of photogrammetric data*, pages 149–155, 1987.
 - [39] D. Freedman and T. Zhang. Active contours for tracking distributions. *IEEE Transactions on Image Processing*, 13(4):518– 526, April 2004.
 - [40] G. Gilboa and S. Osher. Nonlocal linear image regularization and supervised segmentation. *Multiscale Modeling and Simulation*, 6(2):595–630, 2007.
 - [41] E. De Giorgi and L. Ambrosio. New Functionals in the Calculus of Variations. *Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur., Atti. Accad.*, 82(2):199–210, 1988.
 - [42] E. Giusti. *Minimal Surfaces and Functions of Bounded Variation*. Birkhauser, Basel, 1985.
 - [43] T. Goldstein and S. Osher. The Split Bregman Method for L1 Regularized Problems, UCLA CAM Report 08-29, 2008.

-
- [44] L. Grady and G. Funka-lea. Multi-label image segmentation for medical applications based on graph-theoretic electrical potentials. In *Proceedings of the European Conference on Computer Vision*, pages 230–245. Springer, 2004.
 - [45] A. Gray. *Modern Differential Geometry of Curves and Surfaces with Mathematica*. CRC Press, Inc., Boca Raton, FL, USA, 1996.
 - [46] D.M. Greig, B.T. Porteous, and A.H. Seheult. Exact maximum a posteriori estimation for binary images. 1987.
 - [47] R.M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6):610–621, 1973.
 - [48] A. Herbulot, S. Jehan-Besson, M. Barlaud, and G. Aubert. Shape Gradient For Multi-Modal Image Segmentation Using Joint Intensity Distributions. In *International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, 2004.
 - [49] A. Herbulot, S. Jehan-Besson, S. Duffner, M. Barlaud, and G. Aubert. Segmentation of vectorial image features using shape gradients and information measures. *J. Math. Imaging Vis.*, 25(3):365–386, 2006.
 - [50] G. Hermosillo and O.D. Faugeras. Dense image matching with global and local statistical criteria: A variational approach. pages 73–78, 2001.
 - [51] H. Ishikawa and D. Geiger. Segmentation by grouping junctions. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 125–131, 1998.
 - [52] A.K. Jain and F. Farrokhnia. Unsupervised Texture Segmentation Using Gabor Filters. In *IEEE International Conference on Systems, Man and Cybernetics*, pages 14–19, 1990.
 - [53] S. Jehan-Besson, M. Barlaud, and G. Aubert. DREAM2S: Deformable Regions driven by an Eulerian Accurate Minimization Method for Image and Video Segmentation. *International Journal of Computer Vision*, 53(1):45–70, 2003.
 - [54] S. Jehan-Besson, M. Barlaud, G. Aubert, and O. Faugeras. Shape gradients for histogram segmentation using active contours. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, page 408, Washington, DC, USA, 2003. IEEE Computer Society.
 - [55] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active Contour Models. *International Journal of Computer Vision*, 1(4):321–331, 1987.
 - [56] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A.J. Yezzi. Conformal Curvature Flows: From Phase Transitions to Active Vision. In *Archive for Rational Mechanics and Analysis*, volume 134, pages 275–301, 1996.
 - [57] E. Kreyszig. *Differential Geometry*. Paperback, 1991.

-
- [58] A. Kundu and S.K. Mitra. A new algorithm for image edge extraction using a statistical classifier approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):569–577, 1987.
 - [59] V. Kwatra, A. Schödl, I. Essa, G. Turk, and A. Bobick. Graphcut textures: Image and video synthesis using graph cuts. *Proceedings of the Conference on Computer graphics and interactive techniques, SIGGRAPH*, 22(3):277–286, July 2003.
 - [60] T.S. Lee. Image representation using 2d gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):959–971, 1996.
 - [61] L. Liang, C. Liu, Y.Q. Xu, B. Guo, and H.Y. Shum. Real-time texture synthesis by patch-based sampling. *ACM Trans. Graph.*, 20(3):127–150, 2001.
 - [62] J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *International Journal of Computer Vision*, 43(1):7–27, 2001.
 - [63] S.G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, July 1989.
 - [64] S. Marcelja. Mathematical description of the response of simple cortical cells. *J. Optical Soc. Am.*, 70:1297–1300, 1980.
 - [65] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 416–423, July 2001.
 - [66] U. Massari and I. Tamanini. Regularity properties of optimal segmentations. *J.Reine Angew. Math*, 420:61–84, 1991.
 - [67] B. Merriman, J. K. Bence, and S. J. Osher. Motion of multiple junctions: a level set approach. *J. Comput. Phys.*, 112(2):334–363, 1994.
 - [68] J.M. Morel and S. Solimini. Segmentation of Images by Variational Methods: A Constructive Approach. *Rev. Math. Univ. Complut. Madrid*, 1:165–182, 1988.
 - [69] B. Mory and R. Ardon. Fuzzy region competition: A convex two-phase segmentation framework. In *SSVM*, pages 214–226, 2007.
 - [70] B. Mory, R. Ardon, and J.P. Thiran. Variational segmentation using fuzzy region competition and local non-parametric probability density functions. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2007.
 - [71] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.
 - [72] M. Nixon and A.S Aquado. *Feature Extraction and Image Processing*. 2002.

-
- [73] S. Osher and J.A. Sethian. Fronts Propagating with Curvature-Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations. *Journal of Computational Physics*, 79(1)(12-49), 1988.
 - [74] N. Paragios and R. Deriche. Geodesic Active Regions: A New Paradigm to Deal with Frame Partition Problems in Computer Vision. *Journal of Visual Communication and Image Representation*, 13(1-2):249–268, 2002.
 - [75] N. Paragios and R. Deriche. Geodesic Active Regions and Level Set Methods for Supervised Texture Segmentation. *International Journal of Computer Vision*, 46(3): 223–247, 2002.
 - [76] N. Paragios and R. Deriche. Coupled geodesic active regions for image segmentation: A level set approach. In *Proceedings of the European Conference on Computer Vision*, pages 224–240, London, UK, 2000. Springer-Verlag.
 - [77] E. Parzen. On the estimation of a probability density function and the mode. *Annals of Mathematical Statistics*, 33:1065–1076, 1962.
 - [78] A.P. Pentland. Fractal-based description of natural scenes. *PAMI*, 6(6):661–674, November 1984.
 - [79] J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. Mutual information based registration of medical images: a survey. *Med. Image Anal.*, 3:195–206, 1997.
 - [80] A. Protiere and G. Sapiro. Interactive image segmentation via adaptive weighted distances. *IEEE Transactions on Image Processing*, 16(4):1046–1057, 2007.
 - [81] A.R. Rao and B.G. Schunck. Computing oriented texture fields. *CVGIP: Graph. Models Image Process.*, 53(2):157–185, 1991.
 - [82] R. Ronfard. Region-based strategies for active contour models. *Int. J. Comput. Vision*, 13(2):229–251, 1994.
 - [83] A. Rosenfeld, R.A. Hummel, and S.W. Zucker. Scene labeling by relaxation operations. *IEEE Transactions on Systems Man and Cybernetics*, 6(6):420–433, 1976.
 - [84] M. Rousson and R. Deriche. A Variational Framework for Active and Adaptive Segmentation of Vector Valued Images. In *Workshop on Motion and Video Computing (WMVC)*, pages 56–61, 2002.
 - [85] M. Rousson, T. Brox, and R. Deriche. Active unsupervised texture segmentation on a diffusion based feature space. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 699–704, 2003.
 - [86] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992.
 - [87] C. Sagiv, N. Sochen, and Y.Y. Zeevi. Integrated active contours for texture segmentation. *IEEE Transactions on Image Processing*, 15(6):1633–1646, June 2006.
 - [88] C. Samson, L. Blanc-Féraud, G. Aubert, and J. Zerubia. A level set model for image classification. *International Journal of Computer Vision*, 40(3):187–197, 2000.

-
- [89] G. Sapiro. Color snakes. *Comput. Vis. Image Underst.*, 68(2):247–253, 1997.
 - [90] J.A. Sethian. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision and Material Sciences*. Cambridge University Press, 1999.
 - [91] C.E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27:379–423 and 623–656, July and October 1948.
 - [92] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:888–905, 2000.
 - [93] N. Sochen, R. Kimmel, and R. Malladi. A General Framework For Low Level Vision, Technical Report LBNL-39243, Physics Department, Berkeley, 1996.
 - [94] N. Sochen, R. Kimmel, and R. Malladi. A General Framework For Low Level Vision. *IEEE Transactions on Image Processing*, 7(3):310 – 318, 1998.
 - [95] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis, and Machine Vision*. Thomson-Engineering, 2007.
 - [96] A.D. Szlam, R.R. Coifman, and M. Maggioni. A general framework for adaptive regularization based on diffusion processes. *Journ. Mach. Learn. Res.*, November 2007.
 - [97] S. Theodoridis and K. Koutroumbas. *Pattern Recognition, Third Edition*. Academic Press, February 2006.
 - [98] M. Unger, T. Pock, and H. Bischof. Continuous globally optimal image segmentation with local constraints. In *Computer Vision Winter Workshop 2008*, 2008.
 - [99] L. Vese and T. Chan. A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model, UCLA CAM Report 01-25, 2001.
 - [100] L.A. Vese and T. Chan. A multiphase level set framework for image segmentation using the mumford and shah model. *International Journal of Computer Vision*, 50 (3):271–293, December 2002.
 - [101] P. Viola and W.M. Wells. Alignment by maximization of mutual information. *Int. J. Comput. Vision*, 24(2):137–154, 1997.
 - [102] G. Vogiatzis, P. Torr, and R. Cipolla. Multi-view stereo via volumetric graph-cuts. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 391–398, Washington, DC, USA, 2005. IEEE Computer Society.
 - [103] U. von Luxburg. A tutorial on spectral clustering. Technical Report 149, Max Planck Institute for Biological Cybernetics, August 2006.
 - [104] Z. Wang and B. C. Vemuri. An affine invariant tensor dissimilarity measure and its applications to tensor-valued image segmentation. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 228–233, 2004.

-
- [105] J. Weickert. A review of nonlinear diffusion filtering. pages 3–28, London, UK, 1997. Springer-Verlag.
 - [106] Z. Wu and R. Leahy. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1101–1113, 1993.
 - [107] A. Yezzi, A. Tsai, and A. Willsky. A Statistical Approach to Snakes for Bimodal and Trimodal Imagery. In *Proceedings of IEEE Computer Society International Conference of Computer Vision*, pages 898–903, 1999.
 - [108] S.X. Yu and J. Shi. Segmentation given partial grouping constraints. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):173–183, 2004.
 - [109] H.K. Zhao, T. Chan, B. Merriman, and S. Osher. A variational level set approach to multiphase motion. *J. Comput. Phys.*, 127(1):179–195, 1996.
 - [110] S. C. Zhu and A. Yuille. Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900, 1996.
 - [111] X. Zhu. Semi-supervised learning literature survey. 2007.

Curriculum Vitae

Name: Nawal Houhou

Degrees: Bachelor of Science in Physics,
Master of Science in Applied Mathematics

Address: Signal Processing Laboratory (LTS5)
Swiss Federal Institute of Technology (EPFL)
CH-1015 Lausanne
Switzerland

Contact numbers: Tel. +41 21 693 56 51
Fax. +41 21 693 76 00
E-mail: nawal.houhou@epfl.ch

Civil status: Single

Date and place of birth: August 28th 1981, Algiers, Algeria

Nationality: Algerian

Education

- Since 2004* **PhD student**
Signal Processing Laboratory (LTS5)
Swiss Federal Institute of Technology (EPFL)
Lausanne, Switzerland
- 2003–2004* **Master of Science in Applied Mathematics in Image Processing, Vision and Learning**
Ecole Normale Supérieure de Cachan, France
- 1999–2003* **Bachelor of Science in Physics**
Université Denis Diderot, ParisVII , France

Experience

- Since 2004* **Teaching and Research Assistant**
Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland
Provided assistance for various courses, 6 semesters.
- June-August 2008* Visit in University of California (UCLA) under supervision of Dr X.Bresson.
- Since 2006* **Supervisor for Master students**
Signal Processing Laboratory (LTS5) Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland
- May-August 2004* **Internship in Geneva Cantonal Hospital, Digital Imaging Unit**
Project: Segmentation of the heart left ventricle using prior knowledge

Relevant Skills

- Areas of expertise:* Image Processing, Variational Methods.
- Programming:* C/C++, ITK, Matlab, Java.

Languages

- Dialectal Arab:* Mother tongue
- French:* Mother tongue
- English:* fluent
- Spanish:* intermediate

Hobbies

tennis, reading, skiing.

Publications

Journal Papers

- S.Gorthi, V. Duay, N. Houhou, M. Bach Cuadra, U. Becker, A. Allal, J-Ph. Thiran **Segmentation of head and neck lymph node regions for radiotherapy planning, using active contour based atlas registration.** IEEE Journal of selected topics in signal processing, 2008.

Conference Papers

- N. Houhou, X. Bresson, A. Szlam, T. Chan and J-Ph. Thiran. **Non-local Semi-Supervised Segmentation based on Continuous Mincut.** Scale Space and Variational Methods in Computer Vision, Voss, Norway, (Accepted) June 2009.
- V. Duay, N. Houhou, S. Gorthi, A. Allal and J-Ph. Thiran. **Hierarchical Image Registration with an Active Contour-Based Atlas Registration Model.** 16th European Signal Processing Conference, Lausanne, Switzerland, 2008.
- N. Houhou, A. Lemkaddem, V. Duay, A. Allal and J-Ph. Thiran. **Shape Prior Based on Statistical MAP for Active Contour Segmentation.** IEEE International Conference on Image Processing (ICIP), San Diego, USA, 12-15 October, 2008.
- N. Houhou, J-Ph. Thiran and X. Bresson. **Fast Texture Segmentation Model based on the Shape Operator and Active Contour.** Computer Vision and Pattern Recognition, Anchorage, USA, 2008.
- V. Duay, X. Bresson, N. Houhou, M. Bach Cuadra and J-Ph. Thiran. **Registration of Multiple Regions Derived from the Optical Flow Model and the Active Contour Framework.** 15th European Signal Processing Conference (EUSIPCO), Poznan, Poland, September, 2007.

