



Mesures statistiques non-paramétriques pour la segmentation d'images et de vidéos et minimisation par contours actifs

Ariane Herbulot

► To cite this version:

Ariane Herbulot. Mesures statistiques non-paramétriques pour la segmentation d'images et de vidéos et minimisation par contours actifs. Interface homme-machine [cs.HC]. Université Nice Sophia Antipolis, 2007. Français. NNT: . tel-00507087

HAL Id: tel-00507087

<https://theses.hal.science/tel-00507087>

Submitted on 29 Jul 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Nice - Sophia Antipolis

École doctorale STIC (Sciences et Technologies de l'Information et de la
Communication)

THÈSE

pour obtenir le titre de

Docteur en Sciences

Mention : Automatique, Traitement du Signal et des Images

présentée par

Ariane HERBULOT

Equipe d'accueil : CReATiVe

Laboratoire I3S - Université de Nice Sophia-Antipolis, CNRS

MESURES STATISTIQUES NON-PARAMÉTRIQUES POUR LA SEGMENTATION D'IMAGES ET DE VIDÉOS ET MINIMISATION PAR CONTOURS ACTIFS

Thèse dirigée par Pr. Michel Barlaud et Pr. Gilles Aubert

soutenue le 10 octobre 2007 devant le jury composé de

Gilles AUBERT	Professeur des Universités (Nice Sophia-Antipolis)	Président
Françoise DIBOS	Professeur des Universités (Paris 13)	Rapporteur
Philippe REFREGIER	Professeur des Universités (Aix-Marseille)	Rapporteur
Ferran MARQUES	Professeur à l'UPC (Barcelone)	Examineur
Patrick PEREZ	Directeur de recherche à l'INRIA (Rennes)	Examineur
Michel BARLAUD	Professeur des Universités (Nice Sophia-Antipolis)	Directeur de thèse
Luc ROBERT	CTO à RealViz (Sophia Antipolis)	Invité

Table des matières

NOTATIONS	IX
1 INTRODUCTION	1
1.1 Quelques méthodes de segmentation	2
1.1.1 Approches contours	3
1.1.2 Approches régions	3
1.2 Objectifs	5
1.3 Structure du manuscrit	6
I THÉORIE	9
2 CONTEXTE DES CONTOURS ACTIFS	11
2.1 Principe des contours actifs	12
2.2 Vers des contours actifs basés régions	13
2.2.1 Approches variationnelles	13
2.2.2 Approches non variationnelles	15
2.2.3 Contours actifs basés régions	16
2.3 Gradient de forme	18
2.3.1 Introduction	18
2.3.2 Outils d'optimisation de domaine	19
2.3.3 Dérivation de termes basés contour et basés régions	21
2.4 Implémentation	26
2.4.1 Approche implicite : les ensembles de niveaux	26
2.4.2 Approche explicite : les modèles paramétriques	28
2.5 Conclusion	30
3 CRITÈRES STATISTIQUES PROVENANT DE LA THÉORIE DE L'INFORMATION	31
3.1 Introduction	31
3.2 Critère général basé sur les distributions	32
3.2.1 Estimation d'une densité de probabilité	32
3.2.2 Critère fonction de la densité de probabilité	34
3.2.3 Dérivation du critère fonction de la pdf	37
3.3 Entropie	39

3.3.1	Pourquoi l'entropie ?	39
3.3.2	Estimation de l'entropie	40
3.3.3	Entropie jointe et entropie conditionnelle	41
3.4	Divergences	41
3.4.1	Information mutuelle	42
3.4.2	Divergence de Kullback-Leibler	42
3.5	Conclusion	43
II	SEGMENTATION D'IMAGES	45
4	INTRODUCTION À LA SEGMENTATION D'IMAGES	47
5	ENTROPIE	49
5.1	Entropie comme critère de segmentation	49
5.2	Minimisation de l'entropie	50
5.3	Résultats expérimentaux	51
5.3.1	Images de textures	53
5.3.2	Images médicales en niveau de gris	55
5.4	Conclusion	56
6	ENTROPIE JOINTE ET INFORMATION MUTUELLE	61
6.1	Entropie jointe comme critère de segmentation	61
6.1.1	Minimisation de l'entropie jointe	62
6.1.2	Résultats expérimentaux	62
6.2	Information mutuelle	63
6.2.1	Minimisation du critère	67
6.2.2	Résultats expérimentaux	67
6.3	Conclusion	68
7	DIVERGENCE DE KULLBACK-LEIBLER	71
7.1	Critère basé sur une distance entre distributions	71
7.1.1	Définition de la distance entre distributions	71
7.1.2	Vers l'équation d'évolution du contour actif	72
7.2	Résultats expérimentaux	75
7.3	Conclusion	80
8	ENTROPIE ET A PRIORI DE DISTANCE : EXEMPLE AVEC LES WATERSHEDS	87
8.1	Définition de la contrainte géométrique	87
8.1.1	Définition de la distance à un contour	87
8.1.2	Critère d'a priori de distance	88
8.1.3	Dérivation du terme d'a priori géométrique	89
8.2	Entropie et a priori de distance	90
8.3	Application : distance provenant de l'algorithme de la ligne de partage des eaux	90

8.3.1	Utilisation de l'algorithme de la ligne de partage des eaux pour définir une distance	90
8.3.2	Segmentation contrainte par les watersheds	91
8.3.3	Résultats expérimentaux	92
8.4	Conclusion	93
III	SEGMENTATION DE SÉQUENCES VIDÉO	97
9	INTRODUCTION À LA SEGMENTATION DE VIDÉOS	99
9.1	Détection de mouvement	99
9.2	Estimation du mouvement	100
9.3	Présentation des travaux sur la segmentation de séquences vidéo	101
10	SEGMENTATION DE DONNÉES VECTORIELLES : ENTROPIE DU FLOT OPTIQUE	103
10.1	L'entropie pour segmenter le mouvement	103
10.1.1	Entropie	103
10.1.2	Entropie conditionnelle	104
10.2	Résultats sur des images synthétiques	105
10.3	Résultats sur des séquences réelles	106
10.3.1	Estimation du mouvement par flot optique	106
10.3.2	Résultats expérimentaux	111
10.4	Conclusion	113
11	SEGMENTATION SPATIO-TEMPORELLE EN MINIMISANT L'ENTROPIE JOINTE DU RÉSIDUEL ET DE L'INTENSITÉ	117
11.1	Position du problème	117
11.2	Energie proposée	118
11.2.1	Energie temporelle	119
11.2.2	Nécessité de considérer l'information spatiale	121
11.2.3	Combiner les énergies spatiale et temporelle	123
11.3	Vers l'équation d'évolution du contour actif	124
11.4	Résultats expérimentaux	126
11.4.1	Conditions d'expérimentations	126
11.4.2	Comparaisons avec énergie spatiale, temporelle et spatio-temporelle	127
11.4.3	Approche non-paramétrique comparée à l'approche paramétrique	130
11.4.4	Suivi d'objet et estimation du mouvement par morceaux	135
11.5	Conclusion	136
IV	CONCLUSION ET PERSPECTIVES	141
12	CONCLUSION	143

V	ANNEXES	147
A	DISTANCE ENTRE FONCTIONS DE RÉPARTITION	149
A.1	Définition du critère	149
A.2	Dérivation du critère par les gradients de forme	150
B	SEGMENTATION PAR WATERSHEDS	151
B.1	Définition de l'algorithme de la ligne des partage des eaux	151
B.2	Diminution du nombre de régions	153
B.3	Utilisation de cette segmentation pour la méthode des contours actifs	153
C	COMPENSATION DE LA CAMÉRA	155
C.1	Estimation du mouvement de la caméra	155
C.1.1	Mise en correspondance de blocs ou <i>block matching</i>	155
C.1.2	Estimation robuste du modèle de mouvement de la caméra	157
C.2	Compensation du mouvement de la caméra	158
	BIBLIOGRAPHIE	161
	PUBLICATIONS	175

Table des figures

1.1	Diverses applications pour l'image.	1
1.2	Images représentant un paysage et une chèvre	2
1.3	Détection de contours sur l'image Chevre	3
1.4	Détection de régions par seuillage de l'intensité	4
1.5	Algorithme de la ligne de partage des eaux	4
1.6	Algorithme des k -means, partition en 4 groupes.	5
2.1	Evolution du contour actif $\Gamma(\tau)$ avec la vitesse $\mathbf{v}$ vers l'objet d'intérêt	12
2.2	Illustration de la fonction distance signée u	27
2.3	Construction d'une B-spline cubique	29
3.1	Illustration de la méthode de Parzen	33
3.2	Influence de la largeur de bande, h de 0.1	34
3.3	Influence de la largeur de bande, h de 1	35
3.4	Influence de la largeur de bande, h calculé par méthode <i>plug-in</i>	35
3.5	Relations entre entropie, entropie conditionnelle et information mutuelle. . . .	42
5.1	Image composée de textures et histogrammes des textures	53
5.2	Comparaison d'une segmentation en utilisant des distributions paramétriques et non-paramétriques	55
5.3	Image composée de textures et histogrammes des textures	55
5.4	Comparaison d'une segmentation en utilisant des distributions paramétriques et non-paramétriques	56
5.5	Evolution de la segmentation et des histogrammes en minimisant l'entropie . . .	58
5.6	Comparaison des histogrammes et des gaussiennes estimées	59
6.1	Evolution de la segmentation et des histogrammes en minimisant l'entropie jointe	64
6.2	Segmentation de la fleur en utilisant l'entropie jointe (canaux couleurs Y et U)	65
6.3	Segmentation du rhinocéros en utilisant l'entropie jointe (canaux couleurs Y et U)	65
6.4	Segmentation de l'écureuil en utilisant l'entropie jointe (canaux couleurs Y et U)	65
6.5	Evolution de la segmentation en minimisant l'information mutuelle	68

6.6	Evolution de la segmentation en minimisant l'information mutuelle	69
6.7	Segmentation au cours de la séquence vidéo	70
7.1	Segmentation de référence et <i>pdf</i> de référence de l'objet et du fond	76
7.2	Evolution de la segmentation	77
7.3	Evolution de la <i>pdf</i> de l'objet	78
7.4	Evolution de la <i>pdf</i> du fond	79
7.5	Comparaison des <i>pdf</i> finales du fond et de l'objet et des <i>pdf</i> de référence . .	80
7.6	Segmentation de référence et <i>pdf</i> de référence de l'objet et du fond	81
7.7	Evolution de la segmentation sur la séquence Phone	82
7.8	Evolution de la <i>pdf</i> de l'objet	83
7.9	Evolution de la <i>pdf</i> du fond	84
7.10	Comparaison des <i>pdf</i> finales du fond et de l'objet et des <i>pdf</i> de référence . .	85
8.1	Image et squelette intérieur et extérieur de cette image.	88
8.2	Illustration de la notion de distance signée (à gauche) et de la distance non-signée (à droite) à un contour de référence Γ^{ref}	88
8.3	Définition des vecteurs normalisés intérieurs du contour Γ et du contour Γ^{ref} . .	89
8.4	Carte de distances sur une image de la séquence Mother and daughter (en rouge sont représentées les lignes de partage des eaux).	91
8.5	Image issue de la séquence Foreman et carte de distances résultant d'une partition de l'image en 30 régions par l'algorithme des watersheds.	93
8.6	Evolution du contour actif sans la contrainte géométrique (colonne de gauche) et avec la contrainte géométrique (colonne de droite).	94
8.7	Evolution du contour actif sans la contrainte géométrique (colonne de gauche) et avec la contrainte géométrique (colonne de droite).	95
10.1	Rectangle se déplaçant : Segmentation	106
10.2	Rectangles se déplaçant : Segmentation	107
10.3	Rectangles se déplaçant : Evolution des histogrammes	108
10.4	Disque s'élargissant : Segmentation	109
10.5	Segmentation en utilisant la longueur du vecteur.	112
10.6	Segmentation en utilisant les coordonnées du vecteur (u, v)	113
10.7	Evolution des histogrammes en utilisant les coordonnées du vecteur (u, v) . .	114
10.8	Flot optique sur la séquence Taxi	115
10.9	Segmentation de la séquence Taxi en utilisant la longueur du vecteur (à gauche) et les coordonnées du vecteur (u, v) (à droite).	115
11.1	Comparaison des <i>pdf</i> du résiduel et de l'intensité de l'objet avec ou sans partie du fond	122
11.2	Comparaisons des critères temporel, spatial et spatio-temporel pour la segmentation d'objets synthétiques sur des fonds synthétiques.	128
11.3	Segmentation de l'arbre sur la séquence ' Flowers and garden '.	129
11.4	Segmentation d'un joueur sur la séquence ' Soccer '.	130
11.5	Segmentation d'un joueur sur la séquence ' Football '.	131

11.6	Comparaison du terme temporel avec une méthode classique paramétrique de l'image 50 de la séquence Marbled blocks	133
11.7	Segmentation basée mouvement sur la séquence Marbled blocks entre les images 50 et 200.	134
11.8	Comparaison d'une approche paramétrique et d'une non-paramétrique sur la séquence City	135
11.9	Segmentation de l'image 237 de la séquence Flowers and garden en supposant des distributions paramétriques.	136
11.10	Segmentation de l'image 72 de la séquence Football en supposant des distributions paramétriques.	137
11.11	Suivi d'objet sur la séquence Soccer : comparaison entre l'approche du mouvement global (colonne de gauche) et l'approche du mouvement local (colonne de droite) sur les images 164, 172, 182, 192 et 202. La segmentation est plus précise avec l'approche locale (les principales différences sont entourées d'un cercle).	139
11.12	Suivi d'objet sur la séquence Football : comparaison entre l'approche du mouvement global (colonne de gauche) et l'approche du mouvement local (colonne de droite) sur les images 73, 77, 81, 85 et 89. La segmentation est plus précise avec l'approche locale (les principales différences sont entourées d'un cercle).	140
B.1	Ligne de partage des eaux	152
B.2	Image et résultat de l'algorithme de la ligne de partage des eaux.	152
B.3	Illustration de l'arbre de partition binaire.	154
B.4	Fusion de la sur-segmentation.	154
C.1	Estimation du mouvement par méthode <i>backward</i>	156
C.2	Estimation du mouvement par méthode <i>forward</i>	156
C.3	Différence de l'intensité lumineuse entre I_{n-1} et I_n sans compensation de la caméra (à gauche) et avec compensation (à droite).	158
C.4	Flot optique sans (à gauche) et avec (à droite) compensation de la caméra.	159

Notations et abréviations

NOTATIONS PRINCIPALES

- Les vecteurs sont représentés par des caractères gras : par exemple le vecteur $\mathbf{x}$.
- Les matrices sont représentées par des majuscules : par exemple la matrice M .
- La transposée d'un vecteur $\mathbf{v}$ se note $\mathbf{v}^T$, la transposée d'une matrice M se note M^T .
- Les images sont représentées par des fonctions continues $\mathbf{I} : \Omega_I \rightarrow \mathbb{R}^m$ avec Ω_I un sous-ensemble de $\mathbb{R}^2$ représentant le support de l'image. On note $\mathbf{I}(\mathbf{x})$ l'intensité lumineuse du pixel $\mathbf{x} \in \Omega_I$.
- Pour des images couleur, $m = 3$ et l'intensité lumineuse est représentée par une fonction $\mathbf{I} : \Omega_I \rightarrow \mathbb{R}^3$ où $\mathbf{I} = [I^1, I^2, I^3]^T$. Pour des images en niveaux de gris, $m = 1$ et l'intensité lumineuse de l'image est notée I .
- Pour une séquence d'images contenant N images, l'intensité lumineuse de l'image $n \in [1, N]$ est notée $\mathbf{I}_n : \Omega_I \rightarrow \mathbb{R}^m$.
- Soit $\Gamma(p) : [a, b] \rightarrow \mathbb{R}^2$ une courbe paramétrée par p . Le contour est orienté dans le sens inverse des aiguilles d'une montre. Une des paramétrisations possible du contour est l'abscisse curviligne, notée s et définie de façon à ce que $|\Gamma'(s)| = 1$. Dans ce cas, le contour $\Gamma(s)$ est défini sur $[0, L]$ avec L la longueur du contour. Nous notons $\mathbf{T}$ la tangente unitaire au contour, $\mathbf{N}$ la normale unitaire intérieure au contour et κ la courbure du contour.
- La région comprise à l'intérieur du contour Γ est notée Ω_{in} , la région comprise à l'extérieur est notée Ω_{out} .
- Nous noterons MI l'information mutuelle, même si la notation habituelle consiste à la noter I , pour éviter la confusion avec l'image I .

ABRÉVIATIONS ET ACRONYMES

- EDP : Equation aux dérivées partielles.
- I3S : Informatique Signaux et Systèmes de Sophia Antipolis.
- MPEG : *Moving Picture Expert Group* - Groupe d'experts chargés de mettre au point des formats de compression vidéo.
- SIMILAR : *The European research taskforce creating human-machine interfaces SIMILAR to human-human communication - Network of Excellence funded by the IST 6th framework of the European Commission* - Réseau d'excellence européen du 6ème programme-cadre sur les interfaces multimodales.

ANGLICISMES

- *block-matching* : mise en correspondance de blocs.
- *outlier* : point ou échantillon qui se situe loin des autres dans l'espace des paramètres, mesure aberrante.
- *snake* : premier modèle de contour actif introduit par Kass *et al.* [[KWT88](#)].
- *watersheds* : algorithme de la ligne de partage des eaux.
- *pdf* : fonction de densité de probabilité.

Introduction

Le traitement d'images voit ses débuts dans les années 1920 dans la transmission de données par câble mais ne connaît de vrai essor que dans les années 1960 avec le développement des ordinateurs. Au départ, les techniques de traitement d'images sont essentiellement des méthodes de restauration et de compression d'images. Puis se développent avec les progrès de l'informatique, des techniques de détection de primitives (contours, point d'intérêt, lignes d'intérêt ...) et de nombreux autres traitements dans les domaines aussi variés que le médical, la télévision, l'imagerie satellitaire, le multimédia (voir la figure 1.1). C'est dans les années 2000 que l'image numérique et par conséquent le traitement d'images devient omni-présent. Que cela soit sur internet, au cinéma, à la télévision, sur les téléphones, dans le domaine médical, l'image est partout. Aujourd'hui il ne s'agit plus uniquement de traiter les images pour les améliorer mais aussi de les comprendre et de les interpréter. C'est dans ce contexte que la reconnaissance d'objets dans les images devient un sujet de recherche important. Et pour reconnaître des objets afin d'interpréter les images, il faut souvent au préalable les segmenter, c'est-à-dire séparer les objets d'intérêt du fond de l'image.

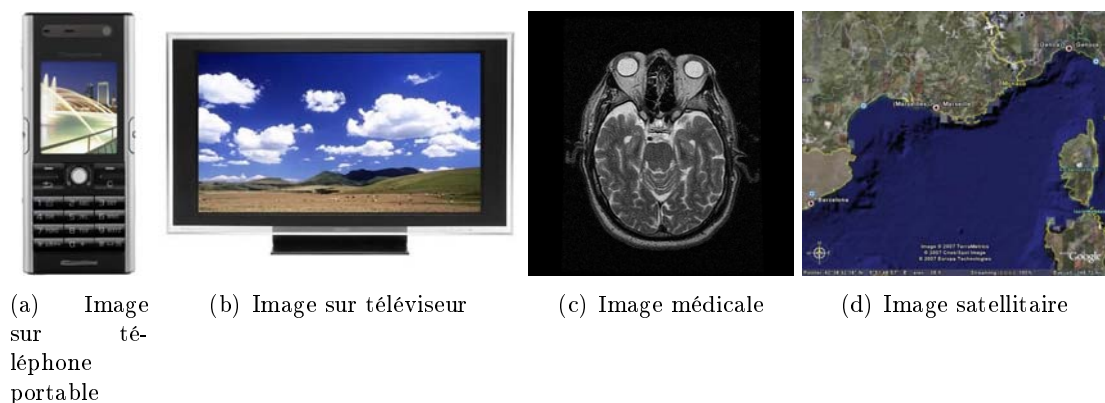


FIG. 1.1 – Diverses applications pour l'image.

1.1 QUELQUES MÉTHODES DE SEGMENTATION

Comment déterminer les objets d'intérêt d'une image ? C'est tout le problème de la segmentation d'images qui est un problème mal posé. La segmentation d'images numériques consiste à regrouper ensemble des pixels qui possèdent les mêmes caractéristiques. Il faut au préalable déterminer ce qui représente au mieux les objets. Cela peut être une couleur particulière, une forme précise. Mais est-ce suffisant ? Pour l'humain, l'extraction des objets du fond de l'image repose sur un ensemble de critères et sur le fait qu'il reconnaît l'objet qui l'intéresse. Une part importante dans l'extraction des objets repose sur l'acquis de l'humain qui connaît les propriétés des objets présents dans la scène. Il sait (par expérience) par exemple qu'une balle est un objet de forme sphérique ou ovoïde et qu'un arbre possède en général un tronc et des feuilles (ou des épines) et qu'en hiver certains arbres n'ont plus de feuilles mais restent néanmoins des arbres. Comment reproduire cela avec un ordinateur qui ne dispose que de pixels ?

Quel est l'objet d'intérêt sur les deux images de la figure 1.2 ? Pour la première c'est assez dur à déterminer étant donné qu'il s'agit d'un paysage. Certains diraient qu'il s'agit de la mer, d'autres d'arbres, d'autres encore de rochers... Pour la seconde, chacun peut clairement affirmer qu'il s'agit d'une chèvre. Nous allons nous intéresser principalement au deuxième type d'image, à savoir des images où un ou plusieurs objets sont séparables du reste de l'image.



FIG. 1.2 – Images représentant un paysage et une chèvre

Il existe de très nombreuses méthodes de segmentation qui varient fortement de par leur façon d'aborder le problème de la segmentation d'images. Un certain nombre de ces méthodes sont recensées dans [FM81, SK94, Ser06]. Nous allons en présenter quelques-unes. Nous pouvons distinguer les approches contours des approches régions.

1.1.1 Approches contours

Les approches contours se caractérisent par le fait qu'elles ne prennent en compte que des informations sur le contour des objets.

Les premiers modèles de segmentation s'appuient sur des détections de changement rapide d'intensité lumineuse ou de couleur, c'est-à-dire les contours de l'image. L'application de **détecteurs de contours** sous la forme de filtres dérivateurs permet d'obtenir les contours des objets présents dans la scène. Nous pouvons citer les approches se basant sur les différences finies comme l'opérateur de gradient, l'opérateur laplacien, les filtres de Sobel, Prewitt, Roberts ou bien des approches reposant sur des critères d'optimalité comme les filtres de Canny-Deriche. Mais ce genre de techniques est peu exploitable car elles donnent des contours non fermés, bruités ou des contours non détectés (voir Fig.1.3).

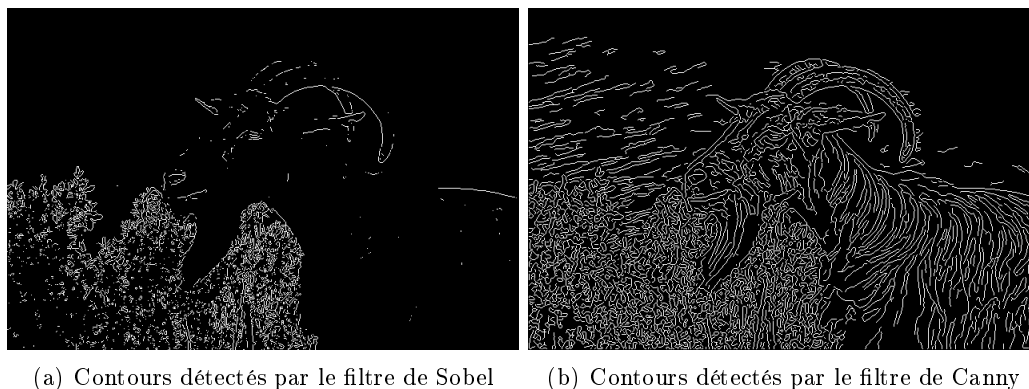


FIG. 1.3 – Détection de contours sur l'image **Chevre**

Mais considérer uniquement les contours pour extraire des objets d'intérêt n'est pas toujours satisfaisant, et prendre en compte les propriétés des régions comprises entre ces contours pourrait améliorer les approches basées uniquement sur les contours.

1.1.2 Approches régions

Nous pouvons tout d'abord noter des méthodes qui effectuent un traitement global sur l'image comme par exemple un **seuillage** [HS92, SSWC88] sur l'intensité de l'image pour déterminer les objets ayant une intensité d'image proche (voir la figure 1.4). Le seuillage peut aussi être déterminé en étudiant l'histogramme de l'intensité de l'image pour repérer les différents modes de la distribution de l'intensité et calculer ainsi des seuils optimaux afin de détecter les différents objets de la scène. Mais cette méthode ne fonctionne que si les objets ont une intensité homogène et de plus, qui soit différente de celle du fond.

Certaines méthodes consistent à définir directement les **caractéristiques des objets** que l'on cherche à segmenter, comme par exemple leur couleur. Les régions obtenues auront donc une couleur homogène. Ce type de méthode est efficace si les objets sont très

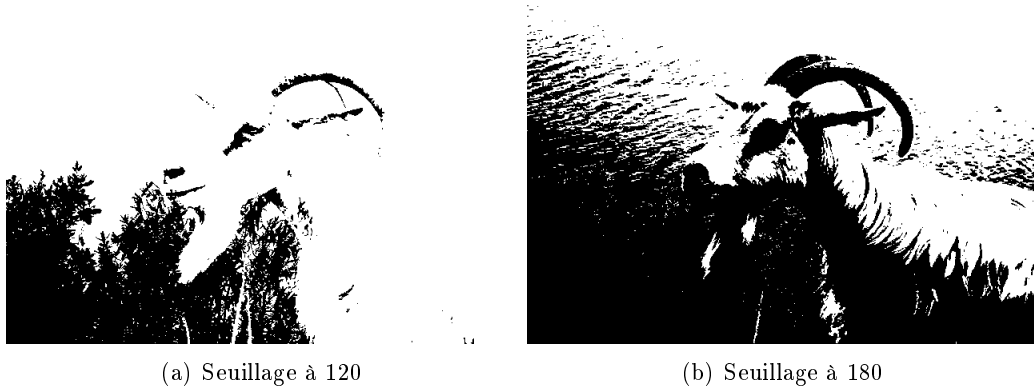


FIG. 1.4 – Détection de régions par seuillage de l'intensité

homogènes, ce qui est rarement le cas.

D'autres approches consistent à faire grandir un ensemble de germes choisis dans l'image. Ce sont les approches par **croissance de régions** [HS92]. Un critère d'homogénéité est choisi afin d'augmenter les régions. Quand plusieurs régions se rejoignent, leurs frontières délimitent les contours des régions finales. Un des principaux inconvénients de cette méthode est que les frontières obtenues ne correspondent pas forcément à des limites précises dans l'image. De plus elle est assez sensible au nombre et au positionnement des germes initiaux.

Des méthodes morphologiques comme l'algorithme de la **ligne de partage des eaux** (LPE, ou *watersheds* en anglais) considèrent le gradient de l'image comme une surface topologique et partent de germes initiaux pour immerger la surface [BM93]. Ainsi, les zones où se rencontrent l'eau des différents bassins versants forment les frontières entre les régions. Ce genre de méthode est très rapide mais fournit un nombre très grand de régions qu'il faudra par la suite fusionner pour obtenir une segmentation correcte des objets de la scène (voir la figure 1.5).

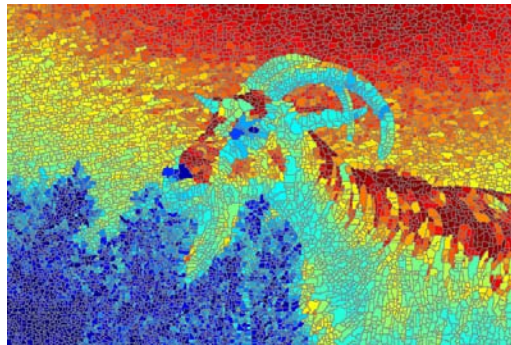


FIG. 1.5 – Algorithme de la ligne de partage des eaux

Il existe des méthodes dites de **division/fusion** (*split and merge* en anglais) [HP77] qui consistent soit à diviser de manière très fine une image et ensuite à fusionner les régions adjacentes selon des critères d'homogénéité, soit à diviser l'image de façon itérative tant que les régions ne sont pas assez homogènes. On utilise souvent des arbres ou des graphes pour effectuer cette étape.

Nous pouvons également indiquer les méthodes de **classification** (plus souvent appelées *clustering* en anglais) qui consistent à regrouper en sous-ensembles les pixels qui possèdent des caractéristiques proches. Citons par exemple les méthodes de k -means [Har75, McQ67] qui consistent à séparer les pixels en k groupes en minimisant une distance entre un pixel donné et le représentant d'un groupe. Ces méthodes fournissent donc une partition de l'image mais ne permettent pas de distinguer quelles classes font parties de l'objet qui nous intéresse, et quelles classes font partie du fond (voir la figure 1.6).



FIG. 1.6 – Algorithme des k -means, partition en 4 groupes.

Toutes les méthodes qui ont été présentées font une partition de l'image en régions, mais ne permettent pas de déterminer directement quelles sont les régions qui font partie de l'objet d'intérêt. C'est pourquoi nous nous intéressons à la méthode des **contours actifs** qui consiste à faire évoluer une courbe initiale vers l'objet d'intérêt. Pour cela il faut déterminer à l'avance quelle caractéristique on cherche à retrouver dans l'objet segmenté afin de pouvoir extraire celui-ci du fond qui ne possède pas ces caractéristiques. Ces caractéristiques peuvent être multiples, il peut s'agir de la couleur, d'une forme précise ou bien d'une combinaison de plusieurs caractéristiques. A partir de ces caractéristiques est élaboré un critère dont le minimum correspondra à la région d'intérêt. Cette fonctionnelle est dérivée afin de parvenir à l'équation d'évolution du contour actif vers l'objet d'intérêt. Nous présentons ces méthodes de façon plus approfondie dans le chapitre 2.

1.2 OBJECTIFS

Notre objectif est de procéder à une segmentation d'objets d'intérêt dans des images et des séquences vidéo. Nous proposons pour cela d'utiliser l'approche par contours actifs puisqu'elle permet d'obtenir directement l'objet d'intérêt. Nous nous intéressons parti-

culièrement aux contours actifs basés régions, c'est-à-dire ceux qui tiennent compte des propriétés des régions à segmenter.

Pour cela nous nous proposons d'élaborer des critères de segmentation qui prennent en compte les vraies distributions des caractéristiques de l'image afin d'approcher au mieux les données de l'image. Nous souhaitons ainsi utiliser des critères non-paramétriques qui permettent d'évaluer au mieux l'information contenue dans les régions. C'est pourquoi nous introduisons des mesures d'information provenant de la théorie de l'information comme critères de segmentation. Etant donné qu'ils reposent sur des distributions réelles et non plus sur des modèles de distribution, ces critères devraient permettre une segmentation plus précise des objets et moins biaisée du fait qu'il n'y a pas de biais apporté par le modèle de distribution.

Nos objectifs sont ainsi d'utiliser ces critères pour segmenter des régions, dont les caractéristiques respectent certaines propriétés liées au critère, dans des images et des vidéos. Dans le cas de critères d'homogénéité comme l'entropie, ces régions seront plus ou moins homogènes, l'entropie permettant une certaine variabilité des caractéristiques considérées, notamment dans le cas de l'utilisation d'une compétition entre la région d'intérêt et le fond. Dans le cas d'utilisation de distance à une segmentation de référence, la région segmentée possédera des caractéristiques proches de celle de référence.

1.3 STRUCTURE DU MANUSCRIT

Dans ce manuscrit, nous présentons l'utilisation des contours actifs en considérant des critères non-paramétriques issus de la théorie de l'information. Nous étudions tout d'abord les différentes mesures d'information que nous utilisons ensuite dans un schéma de contours actifs pour la segmentation d'images et de séquences vidéo.

Ce manuscrit est divisé en trois parties : tout d'abord nous présentons le cadre théorique de nos travaux, puis nous appliquons ce cadre général à la segmentation d'images et enfin à la segmentation de séquences vidéo.

La première partie pose le cadre théorique de nos travaux, du contexte des contours actifs aux critères de segmentation issus de la théorie de l'information.

Chapitre 2 : Contexte des contours actifs. Ce chapitre présente le contexte des contours actifs, des premiers travaux de Kass *et al.* [KWT88] aux méthodes plus récentes basées sur des critères de régions. Nous évoquons les différentes méthodes existantes afin de situer nos travaux par rapport à l'état de l'art. Puis nous présentons la façon de parvenir à l'équation d'évolution du contour actif en partant du critère à minimiser. Enfin les différentes implémentations des contours actifs sont expliquées, ainsi que leurs avantages et inconvénients respectifs.

Chapitre 3 : Critères statistiques provenant de la théorie de l'information. Ce

chapitre met en avant les différents critères statistiques provenant de la théorie de l'information que nous utilisons dans nos travaux. Nous insistons sur le fait que nous ne faisons aucune hypothèse sur les données que nous considérons et de ce fait nous utilisons des critères non-paramétriques. Nous développons le cadre mathématique pour un critère général utilisant des données non-paramétriques puis nous présentons les différents critères qui sont utilisés dans nos travaux, à savoir l'entropie, l'information mutuelle et la divergence de Kullback-Leibler.

La seconde partie est consacrée à l'utilisation des critères définis dans le cadre théorique pour la segmentation d'images fixes.

Chapitre 5 : Entropie. Ce chapitre présente l'utilisation de l'entropie comme critère pour la segmentation d'images fixes. Nous appliquons cet algorithme à la segmentation d'images en niveaux de gris.

Chapitre 6 : Entropie jointe et information mutuelle. Dans ce chapitre nous étendons la notion vue au chapitre précédent en considérant la segmentation de données vectorielles, et plus particulièrement d'images en couleurs. Nous utilisons pour cela des critères basés sur l'entropie jointe et l'information mutuelle.

Chapitre 7 : Divergence de Kullback-Leibler. Dans ce chapitre nous considérons une segmentation de référence et nous cherchons à segmenter une région dont la distribution de l'intensité sera la plus proche possible de celle de référence. Nous cherchons donc à minimiser la divergence de Kullback-Leibler entre la distribution de la région à segmenter et celle de référence.

Chapitre 8 : Entropie et a priori de distance : exemple avec les watersheds. Ce chapitre introduit la notion d'a priori de distance comme critère de segmentation. Nous combinons un critère d'entropie de l'intensité avec une distance à une segmentation obtenue par la méthode des watersheds afin d'avoir une segmentation dont les frontières sont les plus proches possible des contours des objets. Nous combinons donc un critère d'homogénéité et un critère d'a priori de distance.

La troisième partie s'intéresse à la segmentation d'objets mobiles dans des séquences vidéo. Nous présentons deux façons d'utiliser l'entropie afin de segmenter des objets en mouvement.

Chapitre 10 : Segmentation de données vectorielles : entropie du flot optique. Ce chapitre présente la segmentation d'objets mobiles en ayant au préalable estimé le mouvement des objets de la séquence. Nous cherchons à minimiser l'entropie des vecteurs mouvement de chaque point de l'image afin de détecter les objets qui ont un mouvement cohérent.

Chapitre 11 : Segmentation spatio-temporelle en minimisant l'entropie jointe

du résiduel et de l'intensité. Ce chapitre présente une méthode de segmentation d'objets en estimant conjointement le mouvement des objets. Le critère de segmentation tient compte de caractéristiques spatiales et temporelles pour permettre une segmentation des objets en mouvement.

Première partie

Théorie

Chapitre 2

Contexte des contours actifs

Comme nous l'avons vu dans le chapitre 1, notre objectif est de segmenter des images et des vidéos. La segmentation d'images consiste soit à effectuer une partition de l'image en la découpant en zones possédant des caractéristiques proches, soit à extraire une région spécifique de l'image que nous appelons objet d'intérêt. Dans notre étude nous nous intéressons à la recherche d'objets d'intérêt (visage dans une image, objet en mouvement dans une vidéo par exemple). Nous cherchons donc à regrouper ensemble des portions de l'image (ou de la vidéo) possédant les mêmes caractéristiques. Nous nous attachons principalement à définir une fonctionnelle permettant de caractériser au mieux l'objet d'intérêt, celui-ci étant représenté par la zone de l'image qui minimise cette fonctionnelle. Nous nous intéressons donc aux méthodes variationnelles, c'est-à-dire celles qui minimisent un critère. Une grande difficulté repose dans la définition de ce critère. Comment définir les caractéristiques qui nous intéressent et qui vont représenter au mieux l'objet d'intérêt ? De plus, comment minimiser ce critère ?

Les contours actifs apparaissent comme un outil pertinent pour parvenir à implémenter ce problème de minimisation puisqu'ils permettent de faire évoluer un contour vers l'objet d'intérêt. De plus, nous pouvons ajouter un a priori géométrique sur l'objet à segmenter en imposant des contraintes de régularisation sur le contour.

Pour faire évoluer le contour actif vers le minimum de la fonctionnelle, nous devons dériver la fonctionnelle puis en déduire l'équation aux dérivées partielles (EDP) qui fera évoluer le contour vers la région qui correspond, idéalement, au minimum de la fonctionnelle. Notons que toutes les méthodes de contours actifs ne cherchent pas à minimiser une fonctionnelle, certaines introduisent directement l'EDP pour faire évoluer le critère. Nous préciserons cela dans la section 2.2.

Ce chapitre d'état de l'art se décompose de la façon suivante : dans la section 2.1, nous présentons le principe des contours actifs et nous introduisons les notations. La section 2.2 décrit les différents contours actifs, des premiers modèles de contours actifs aux modèles plus récents appelés *contours actifs basés régions*. Ensuite la section 2.3 présente la façon de dériver la fonctionnelle caractérisant la région afin de parvenir à l'équation d'évolution du contour actif. Enfin la section 2.4 propose différentes approches existantes pour implémenter les contours actifs, en particulier deux méthodes : la méthode des ensembles de niveaux et la méthode paramétrique des B-splines.

2.1 PRINCIPE DES CONTOURS ACTIFS

Le principe des contours actifs est de faire évoluer une courbe initiale vers l'objet d'intérêt. L'évolution de cette courbe se fait sous l'action d'une force, qui est généralement déduite de la minimisation d'une fonctionnelle. Notons Γ ce contour actif, défini comme une courbe paramétrée par $p \in [a, b]$ avec $\tau \in [0, T]$ le paramètre d'évolution de la courbe. Γ est défini tel que :

$$\begin{cases} [a, b] \times [0, T] & \rightarrow \mathbb{R}^2 \\ (p, \tau) & \rightarrow \Gamma(p, \tau) = \mathbf{x}(p, \tau) = \begin{pmatrix} x(p, \tau) \\ y(p, \tau) \end{pmatrix} \end{cases}$$

L'évolution de ce contour actif Γ est régie par une équation aux dérivées partielles de la forme suivante :

$$\begin{cases} \frac{\partial \Gamma(p, \tau)}{\partial \tau} = \mathbf{v}(p, \tau) \\ \Gamma(p, 0) = \Gamma_0(p) \end{cases} \quad (2.1)$$

Γ_0 est le contour initial qui peut être défini manuellement par exemple et $\mathbf{v}$ est la vitesse d'évolution de la courbe. La figure 2.1 illustre cette évolution.

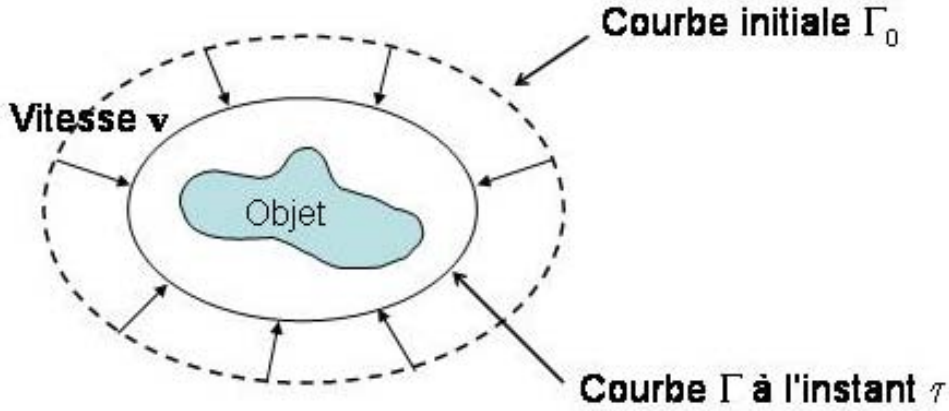


FIG. 2.1 – Evolution du contour actif $\Gamma(\tau)$ avec la vitesse $\mathbf{v}$ vers l'objet d'intérêt

La vitesse d'évolution $\mathbf{v}$ du contour actif a *a priori* une direction quelconque. Décomposons cette vitesse dans le repère de Frenet associé à la courbe. Nous obtenons deux composantes, une composante selon la tangente unitaire $\mathbf{T}$ et une composante selon la normale unitaire intérieure à la courbe $\mathbf{N}$. Cela donne :

$$\frac{\partial \Gamma(p, \tau)}{\partial \tau} = \mathbf{v}(p, \tau) = F_N(p, \tau)\mathbf{N}(p, \tau) + F_T(p, \tau)\mathbf{T}(p, \tau), \quad (2.2)$$

où F_N est l'amplitude de la vitesse dans la direction normale au contour et F_T est l'amplitude de la vitesse dans la direction tangentielle au contour. La composante normale de

la vitesse modifie la géométrie de la courbe tandis que la composante tangentielle de la vitesse modifie sa paramétrisation. L'équation (2.2) peut être simplifiée en considérant que nous nous intéressons uniquement à la déformation de la courbe et non à sa paramétrisation [EC87]. La partie de l'équation qui nous intéresse est donc uniquement la projection sur $\mathbf{N}$ car c'est celle qui régit le déplacement du contour, tandis que la projection sur $\mathbf{T}$ ne permet que de déterminer la paramétrisation du contour. L'équation d'évolution du contour actif s'écrit donc uniquement avec une vitesse normale au contour :

$$\begin{cases} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= \mathbf{v}(p, \tau) = F_N(p, \tau) \mathbf{N}(p, \tau) \\ \Gamma(p, 0) &= \Gamma_0(p) \end{cases} \quad (2.3)$$

Par la suite F_N sera noté F pour simplifier.

Après cette rapide définition du principe des contours actifs, nous présentons maintenant les différents types de contours actifs existant : les contours actifs basés contour et les contours actifs basés régions. Un contour actif basé contour ne tient compte que des caractéristiques locales de l'image, à savoir le long du contour actif. Ce modèle est celui suivi par les premiers modèles de contours actifs, les contours actifs géodésiques et les contours actifs géométriques. Un contour actif basé régions prend en compte les caractéristiques de la région comprise à l'intérieur du contour. C'est au modèle basé régions que nous nous intéressons principalement.

2.2 VERS DES CONTOURS ACTIFS BASÉS RÉGIONS

Dans cette section nous présentons un rapide historique des contours actifs, des premiers *snakes* aux modèles basés régions. Nous distinguons les méthodes variationnelles, *i.e.* celles qui nécessitent la minimisation d'une fonctionnelle des approches qui introduisent directement l'équation d'évolution du contour actif.

2.2.1 Approches variationnelles

2.2.1.1 Contours actifs classiques : les snakes

Les premiers contours actifs ont été introduits par Kass *et al.* [KWT88] en 1988, sous le nom de *snakes*. Considérant une courbe C^2 notée $\Gamma : [a, b] \rightarrow \mathbb{R}^2$, l'équation d'évolution de ce modèle de contours actifs est obtenue en minimisant la fonctionnelle suivante :

$$J(\Gamma) = \alpha \int_a^b |\Gamma'(p)|^2 dp + \beta \int_a^b |\Gamma''(p)|^2 dp + \lambda \int_a^b g(|\nabla I(\Gamma(p))|) dp \quad (2.4)$$

avec α , β et λ des constantes positives.

Les deux premiers termes imposent une contrainte de régularité sur le contour et définissent son élasticité (premier terme) et sa rigidité (second terme) tandis que le troisième terme est un terme d'attache aux données qui va permettre d'attirer le contour actif vers les forts

gradients de l'image. La fonction g est une fonction positive et strictement décroissante. Dans l'article original de Kass *et al.*, la fonction choisie pour g est :

$$g(|\nabla I|) = -|\nabla I|^2$$

Cette première approche de contours actifs a été très largement utilisée mais comporte cependant quelques inconvénients. Tout d'abord, la fonctionnelle (2.4) n'est pas intrinsèque, c'est-à-dire qu'elle dépend de la paramétrisation du contour. D'autre part, la contrainte de régularité est telle qu'elle ne permet de détecter que des objets convexes et qu'elle impose la détection d'un unique objet (les changements de topologie ne sont pas permis). De plus, pour pouvoir converger vers l'objet d'intérêt, le contour initial du contour actif doit être assez proche de l'objet à segmenter. Enfin, la minimisation d'une telle fonctionnelle fait apparaître des termes qui nécessitent le calcul de dérivées d'ordre 4, ce qui pose des problèmes de discrétisation et d'instabilités numériques.

2.2.1.2 Contours actifs géodésiques

La première approche de contours actifs a été améliorée par Caselles *et al.* [CKS97] qui ont introduit les contours actifs géodésiques. La fonctionnelle définie par Kass en (2.4) est modifiée en mettant β à zéro, et de ce fait, ce terme de régularisation jugé trop contraignant est abandonné. De plus, les auteurs démontrent que résoudre le problème des contours actifs classiques avec β à zéro revient à chercher une courbe géodésique dans un espace de Riemann, dans lequel la métrique dépend de l'image. Ils montrent que minimiser (2.4) lorsque $\beta = 0$ revient à minimiser :

$$J(\Gamma) = \int_a^b g(|\nabla I(\Gamma(p))|) |\Gamma'(p)| dp . \quad (2.5)$$

Ce qui peut s'écrire, puisque $|\Gamma'(p)|dp = ds$:

$$J(\Gamma) = \int_0^{L(\Gamma)} g(|\nabla I(\Gamma(s))|) ds, \quad (2.6)$$

avec $L(\Gamma)$ la longueur euclidienne de la courbe Γ et s l'abscisse curviligne de la courbe. La fonctionnelle (2.6) est intrinsèque, *i.e.* qu'elle ne dépend pas de la paramétrisation. L'équation d'évolution suivante est obtenue :

$$\frac{\partial \Gamma}{\partial \tau} = (g(|\nabla I(\Gamma)|) \kappa - \nabla g(|\nabla I(\Gamma)|) \cdot \mathbf{N}) \mathbf{N}, \quad (2.7)$$

avec κ la courbure du contour Γ . Dans [CKS97] est ajoutée une force ballon c , introduite par Cohen [Coh91] :

$$\frac{\partial \Gamma}{\partial \tau} = (g(|\nabla I(\Gamma)|)(c + \kappa) - \nabla g(|\nabla I(\Gamma)|) \cdot \mathbf{N}) \mathbf{N} . \quad (2.8)$$

Si la constante c est positive, le contour va rétrécir, et si elle est négative, le contour va s'élargir, tout en "combattant" la courbure pour permettre de segmenter des formes

concaves et permettre au contour d'évoluer lorsque la courbure est nulle. Cette force rajoute une contrainte sur l'aire de la région. De plus, cette force ballon permet au contour d'éviter de rester "collé" dans des minima locaux de la fonctionnelle, notamment en présence de bruit. Elle permet aussi une plus grande liberté dans le choix du contour initial puisque celui-ci peut être à l'intérieur comme à l'extérieur de l'objet à segmenter (selon le signe choisi pour c).

En suivant cette équation, le contour actif, dit *géodésique* aura un comportement différent selon qu'il se trouve dans une zone homogène ou proche d'un contour. Dans une zone homogène, le gradient de l'image est faible et l'équation se réduit à une évolution en fonction de la courbure. Près d'un contour, le gradient est plus fort et le contour actif évoluera vers le contour de l'objet, qu'il soit à l'intérieur ou à l'extérieur de l'objet. Grâce au dernier terme de l'équation (2.7), le contour actif est attiré aux abords de l'objet et stabilisé sur le contour. Notons que l'équation (2.7) a aussi été introduite dans [KKO⁺95, KKO⁺96].

2.2.2 Approches non variationnelles

D'autres approches de contours actifs ont été proposées en parallèle en introduisant directement une équation d'évolution sans minimiser une fonctionnelle. Il s'agit dans ce cas d'envisager l'évolution de la courbe comme une propagation d'un front d'onde. C'est alors un modèle géométrique et non plus paramétrique. C'est le cas notamment de Caselles *et al.* [CCCD93] et Malladi *et al.* [MSV95, MSV96] qui ont présenté les contours actifs géométriques en s'inspirant des travaux de Osher et Sethian [OS88] sur l'équation de la diffusion de la chaleur. L'équation d'évolution proposée se présente sous la forme suivante :

$$\frac{\partial \Gamma}{\partial \tau} = g(|\nabla I(\Gamma)|) \kappa \mathbf{N} \quad (2.9)$$

avec $g(|\nabla I|)$ une fonction de détection. Elle est définie dans [CCCD93, MSV95] par :

$$g(|\nabla I|) = \frac{1}{1 + |\nabla(G_\sigma * I)|^p}, \quad p \in \{1, 2\}$$

où $G_\sigma * I$ est la convolution de l'image par une gaussienne de variance σ .

Cette fonctionnelle permet la segmentation de l'enveloppe convexe des objets. Pour parvenir à la segmentation d'objets non convexes, il a fallu rajouter un terme semblable à la force ballon introduite par Cohen [Coh91] pour permettre au contour d'évoluer dans un sens ou dans l'autre. L'équation d'évolution proposée par Caselles [CCCD93] est alors :

$$\frac{\partial \Gamma}{\partial \tau} = g(|\nabla I(\Gamma)|)(c + \kappa) \mathbf{N} \quad (2.10)$$

avec c une constante qui détermine la force ballon. Nous retrouvons une équation semblable au premier terme de celle présentée en (2.8) sauf que cette équation ne dérive pas de la minimisation d'une énergie. Nous pouvons remarquer que par rapport aux contours actifs

classiques, nous n'avons conservé que le terme de régularisation du second ordre. De ce fait, la régularisation est moins stricte et son approximation numérique moins délicate. La même force ballon que dans les contours géodésiques est aussi introduite.

Tous ces modèles de contours actifs sont jusqu'à présent basés contours, c'est-à-dire qu'ils ne prennent en compte que des propriétés locales de l'image. Cependant une image est aussi caractérisée par des propriétés plus globales, à savoir les propriétés des régions comprises entre les contours de l'image. C'est pourquoi les contours actifs *basés régions* ont été introduits.

2.2.3 Contours actifs basés régions

Le principe des contours actifs basés régions est donc de prendre en compte des propriétés de l'image sur des régions et non plus localement. La fonctionnelle à minimiser dans le cas de termes régions s'écrit comme une intégrale sur la région et non plus sur le contour. Elle est de la forme :

$$J(\Omega) = \int_{\Omega} k(\mathbf{x}, \Omega) d\mathbf{x} \quad (2.11)$$

où k est appelé descripteur de région. C'est une fonction qui dépend des propriétés de la région Ω . Il est important de noter que souvent les approches région combinent à la fois des termes basés régions et des termes basés contours, principalement pour prendre en compte les caractéristiques des régions et ajouter des contraintes sur le contour.

Les approches basées régions peuvent reposer sur une modélisation statistique des régions à segmenter. Un descripteur de moyenne ou bien de variance sur la région est, par exemple, parfois utilisé.

Les premiers à avoir utilisé des descripteurs de région sont Cohen *et al.* [CBA93] et Ronfard [Ron94]. Cohen *et al.* présentent une méthode de reconstruction de surface en utilisant des contours actifs. Ils incluent un critère contenant des termes basés régions. Ronfard quant à lui propose une méthode de segmentation d'images en deux régions. Il définit la vitesse comme étant proportionnelle à la différence des descripteurs de la région des objets et de ceux de la région du fond.

Dans de nombreux travaux, des critères statistiques sont utilisés. Nous pouvons notamment citer des méthodes développées dans le cadre du MDL (*Minimum Description Length*) [ZY96, GR96, CRB99, FLJ00, RR01, GR04, MRGG06]. Zhu et Yuille [ZY96] proposent un algorithme appelé *compétition de régions*. Leur idée est de présenter un cadre statistique pour combiner les caractéristiques géométriques des contours actifs et les techniques statistiques de croissance de régions. Dans leurs travaux, les descripteurs des régions sont les lois qui régissent les densités de probabilité. L'idée est de supposer connues les lois des densités de probabilité dont les paramètres sont à déterminer. Les paramètres des distributions sont estimés alternativement avec l'optimisation de la partition. Germain *et al.* [GR96] présentent un algorithme basé sur des contours actifs statistiques polygonaux

permettant une bonne localisation des contours dans des images SAR (*Synthetic Aperture Radar*, images issues d'un radar à synthèse d'ouverture). Chesnaud *et al.* [CRB99] utilisent des descripteurs statistiques dans un cadre de segmentation non supervisé. Les caractéristiques statistiques sont estimées conjointement à l'étape de segmentation et cette méthode permet la segmentation d'images bruitées. Figueiredo *et al.* [FLJ00] proposent une méthode permettant de déterminer automatiquement l'ordre du modèle utilisé pour décrire le contour tandis que Ruch *et al.* [RR01] généralisent cette approche à des contours polygonaux et permettent ainsi une bonne régularisation du contour. Galland *et al.* [GR04] présentent un modèle de segmentation adapté à la segmentation de plusieurs régions d'intensité différente tandis que Martin *et al.* [MRGG06] proposent une méthode de segmentation où aucun paramètre n'est à régler et qui peut s'appliquer pour n'importe quelle distribution d'intensité.

Nous pouvons aussi citer les méthodes utilisant les *graph cuts* (méthode de coupe minimale/flot maximal dans un graphe) [GPS89] qui permettent d'effectuer une minimisation d'énergie sans tomber dans un minimum local, et ce en un temps polynômial. Ces méthodes connaissent un fort succès actuellement et sont utilisées dans de nombreuses applications de traitement d'images, et notamment en segmentation d'images [BJ01, XBA03, BRB⁺04, ZK04, RKB04, FZ05].

Chan et Vese [CV01] étendent quant à eux les résultats de Mumford et Shah [MS89] en utilisant la moyenne comme descripteur pour la segmentation en deux régions. Ils développent ces travaux dans le cadre des ensembles de niveaux (que nous verrons plus en détails dans la section 2.4). Dans les travaux de Yezzi *et al.* [YTW99], le descripteur utilisé est aussi la moyenne des régions mais la fonctionnelle est différente puisque c'est la distance quadratique entre les moyennes qui est maximisée. Paragios *et al.* [PD00] proposent d'utiliser des *régions actives géodésiques* dans un cadre supervisé où les paramètres des distributions sont déterminés au préalable. Jehan-Besson *et al.* [JBBA03] utilisent aussi des descripteurs de moyenne mais présentent surtout un cadre général de différenciation des fonctionnelles à l'aide d'outils inspirés de l'optimisation de domaines.

Les contours actifs basés régions sont aussi largement utilisés dans l'imagerie médicale. Debreuve *et al.* [DBAD01] introduisent un critère spatio-temporel pour la segmentation de séquences cardiaques. Montagnat *et al.* [MDM99] utilisent des modèles déformables 3D pour la segmentation de volumes.

Du point de vue de l'implémentation, Precioso *et al.* [PBBU05] étendent le modèle des splines d'interpolation pour en proposer une implémentation paramétrique pour les contours actifs. Ils présentent les splines d'approximation qui permettent une robustesse au bruit intéressante.

Les histogrammes peuvent être utilisés dans des critères de segmentation. Jehan-Besson *et al.* [ABFJB03] et Freedman *et al.* [FZ04] proposent d'utiliser des histogrammes comme descripteurs des régions. Des histogrammes connus sont pris en compte afin de servir de

référence pour la segmentation. Kim *et al.* [KFY⁺05] proposent quant à eux d'utiliser des distributions non-paramétriques afin de minimiser l'information mutuelle entre l'intensité de l'image et un label définissant la position du pixel par rapport au contour.

Nous avons présenté les différentes approches pour définir un critère de segmentation pour les contours actifs, de l'approche purement "contours" aux approches plus récentes qui intègrent des termes de régions dans les fonctionnelles à minimiser. Les approches contour ne prenant en compte que de l'information locale sont assez sensibles au bruit et lors de la minimisation de la fonctionnelle, il n'est pas rare de tomber dans un minimum local. Quant aux approches régions, la plupart des méthodes sont efficaces dans des cas simples où l'objet à segmenter est facilement séparable du fond, de par sa couleur par exemple, mais la segmentation est plus difficile lorsque les objets à segmenter ne respectent pas les hypothèses faites sur leurs distributions.

Maintenant que nous avons introduit les différentes approches de contours actifs, voyons comment à partir de la définition d'une fonctionnelle qui caractérise la région, nous pouvons parvenir à l'équation d'évolution du contour actif. Nous introduisons pour cela la notion de gradient de forme.

2.3 GRADIENT DE FORME

Après avoir introduit les contours actifs basés régions et leur historique, nous allons voir comment à partir d'une fonctionnelle nous pouvons obtenir l'équation d'évolution du contour actif. Pour ce faire nous utilisons une méthode variationnelle décrite précisément dans [JBBA03], à savoir la notion de *gradients de forme*, afin de tenir compte dans la dérivation des variations de la région.

2.3.1 Introduction

Nous nous intéressons uniquement aux approches variationnelles, c'est-à-dire celles qui nécessitent la minimisation d'un critère. L'objet d'intérêt est défini par un domaine qui minimise une certaine fonctionnelle. Pour effectuer cette minimisation, il est donc nécessaire de dériver la fonctionnelle. L'introduction d'un schéma dynamique permet de dériver la fonctionnelle par rapport à un paramètre d'évolution du domaine. Nous choisissons d'utiliser des outils provenant de l'optimisation de domaine pour effectuer cette dérivation mais elle peut aussi s'effectuer par le calcul des variations. L'évolution du contour actif est ensuite déduite de ce calcul de dérivée.

Nous définissons un terme basé contour comme étant une intégrale sur le contour Γ d'un descripteur de contour $k(s, \Gamma)$:

$$J(\Gamma) = \int_{\Gamma} k(s, \Gamma) ds$$

Et de même nous définissons un terme basé région comme étant une intégrale sur la région

d'un descripteur de région $k(\mathbf{x}, \Omega)$:

$$J(\Omega) = \int_{\Omega} k(\mathbf{x}, \Omega) d\mathbf{x}$$

Dans cette section nous décrivons comment en utilisant les gradients de forme nous pouvons calculer la dérivée d'une fonctionnelle ayant des termes basés contour et des termes basés régions. Nous pouvons déjà remarquer que certains termes basés régions (ou contour) dépendent doublement de la région (ou du contour) puisque ce sont des intégrales sur la région (ou le contour) d'un descripteur pouvant lui-même dépendre de la région (ou du contour). Nous verrons comment dériver ce genre de termes doublement dépendant de la région. Nous illustrerons enfin cette dérivation à l'aide d'exemples de descripteurs.

2.3.2 Outils d'optimisation de domaine

Nous présentons ici les outils de dérivation qui vont nous permettre de parvenir à l'équation d'évolution du contour actif. En premier lieu, il convient d'introduire un schéma dynamique pour pouvoir calculer la dérivée du domaine par rapport au paramètre d'évolution du contour. Ensuite, nous définissons les notions de dérivées nécessaires au calcul de la dérivée de la fonctionnelle. Ces notions sont largement inspirées de [SZ92, DZ01].

2.3.2.1 Schéma dynamique

L'ensemble des domaines de $\mathbb{R}^n$ n'a pas de structure d'espace vectoriel, il n'est donc pas possible de dériver la fonctionnelle par rapport au domaine Ω (ou Γ pour des termes basés contour). Les techniques de dérivation de domaines consistent à reporter les variations de domaines sur des fonctions appartenant à un espace vectoriel. Une famille de transformations T_τ , $0 \leq \tau < \epsilon$, est alors introduite pour faire évoluer la région Ω (et le contour Γ dans le cas de termes basés contour) en fonction de τ :

$$T_\tau : \begin{cases} \Omega \rightarrow \Omega(\tau) & \text{avec } T_0(\Omega) = \Omega \\ \Gamma \rightarrow \Gamma(\tau) & \text{avec } T_0(\Gamma) = \Gamma \end{cases} \quad (2.12)$$

avec pour tout $\mathbf{x} \in \mathbb{R}^n$, $T_\tau(\mathbf{x}) \in C^1([0, \epsilon])$.

Nous définissons le champs des vecteurs vitesse $\mathbf{V}$ correspondant à T_τ de la façon suivante :

$$\mathbf{V}(\tau, \mathbf{x}) = \frac{\partial T_\tau}{\partial \tau}(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega \quad \forall 0 \leq \tau < \epsilon \quad (2.13)$$

Cette vitesse est considérée comme la direction de déformation.

Nous obtenons ainsi une suite de fonctionnelles dépendant de τ :

$$J(\Omega(\tau)) = \int_{\Omega(\tau)} k(\mathbf{x}, \Omega(\tau)) d\mathbf{x} \quad (2.14)$$

pour une fonctionnelle basée régions, et :

$$J(\Gamma(\tau)) = \int_{\Gamma(\tau)} k(s, \Gamma(\tau)) ds \quad (2.15)$$

pour une fonctionnelle basée contour.

Le schéma dynamique est donc ici directement introduit dans les fonctionnelles à minimiser.

Le contour actif évolue avec le paramètre d'évolution τ dans la fonctionnelle à minimiser.

2.3.2.2 Théorème de dérivation

Nous pouvons écrire un développement limité au premier ordre de la transformation T_τ , ce qui donne :

$$\begin{aligned} T_\tau(\Omega) &= T_0(\Omega) + \tau \frac{\partial T_0}{\partial \tau}(\Omega) \\ &= \Omega + \tau \frac{\partial T_0}{\partial \tau}(\Omega) \end{aligned} \quad (2.16)$$

Or nous avons :

$$\begin{cases} \mathbf{V}(\tau, \Omega(\tau)) &= \frac{\partial}{\partial \tau}(\Omega(\tau)) = \frac{\partial T_\tau}{\partial \tau}(\Omega) \\ \mathbf{V}(0, \Omega(0)) &= \frac{\partial T_0}{\partial \tau}(\Omega) = \mathbf{V} \end{cases} \quad (2.17)$$

Donc :

$$T_\tau(\Omega) = \Omega(\tau) = \Omega + \tau \mathbf{V} \quad (2.18)$$

Notons que la dérivée de domaine de $k(\mathbf{x}, \Omega)$ dans la direction $\mathbf{V}$, notée $k'(\mathbf{x}, \Omega, \mathbf{V})$ est donnée par :

$$k'(\mathbf{x}, \Omega, \mathbf{V}) = \lim_{\tau \rightarrow 0} \frac{k(\mathbf{x}, \Omega(\tau)) - k(\mathbf{x}, \Omega)}{\tau}$$

Le théorème suivant permet de calculer la dérivée eulérienne d'une fonctionnelle en fonction d'autres dérivées.

Théorème 2.1. *La dérivée eulérienne de la fonctionnelle $J(\Omega) = \int_{\Omega} k(\mathbf{x}, \Omega) d\mathbf{x}$ dans la direction du champ de vecteurs $\mathbf{V}$ est donnée par :*

$$dJ(\Omega, \mathbf{V}) = \int_{\Omega} k'(\mathbf{x}, \Omega, \mathbf{V}) d\mathbf{x} - \int_{\Gamma} k(s, \Omega) (\mathbf{V} \cdot \mathbf{N}) ds$$

où $k'(\mathbf{x}, \Omega, \mathbf{V})$ est la dérivée de domaine de $k(\mathbf{x}, \Omega)$ dans la direction $\mathbf{V}$, $\mathbf{N}$ est la normale unitaire intérieure à Γ et s son abscisse curviligne.

Ce théorème traduit le principe de conservation pour un domaine que l'on suit dans son mouvement. Le premier terme correspond à la variation de k sur le domaine tandis que le second terme prend en compte le mouvement du domaine.

La démonstration de ce théorème peut être trouvée dans [SZ92, DZ01] et une démonstration plus élémentaire dans [JBBA03].

Ce théorème va nous permettre de calculer les dérivées eulériennes de termes basés régions, en convertissant des intégrales de régions en intégrales de contour, ce qui permet d'obtenir facilement l'équation d'évolution du contour actif.

2.3.3 Dérivation de termes basés contour et basés régions

Nous allons étudier la dérivation d'un terme basé contour puis d'un terme basé régions dans le cas où le descripteur est indépendant de la région et dans le cas où il en dépend.

Dans la suite, nous allons définir des critères contours et régions dont la dérivée dite de forme peut s'écrire comme :

$$dJ(\Omega, \mathbf{V}) = \int_{\Gamma} \alpha \mathbf{V} \cdot \mathbf{N} ds \quad (2.19)$$

$$= \langle \alpha \mathbf{N}, \mathbf{V} \rangle_{L_2} \quad (2.20)$$

avec α une fonction de s . Donc $\alpha \mathbf{N}$ apparaît comme le gradient au sens L_2 du critère. Dans une approche de descente de gradient, il convient de choisir la plus grande pente, ce qui revient à prendre :

$$\mathbf{V} = -\alpha \mathbf{N} \quad (2.21)$$

comme déformation du contour. De cette façon nous obtenons l'équation d'évolution du contour actif qui permet de faire avancer le contour vers la région Ω correspondant au minimum de la fonctionnelle J :

$$\frac{\partial \Gamma(p, \tau)}{\partial \tau} = -\alpha \mathbf{N} \quad (2.22)$$

2.3.3.1 Dérivation d'un terme basé contour

Considérons un terme basé contour que l'on écrit sous la forme :

$$J(\Gamma) = \int_{\Gamma} k(s, \Gamma) ds \quad (2.23)$$

La dérivée d'un tel terme peut être trouvée par exemple dans [SZ92, HR03].

Considérons le cas particulier où $k(s, \Gamma) = k(s)$. La dérivée eulérienne de la fonctionnelle $J(\Gamma) = \int_{\Gamma} k(s) ds$ dans la direction $\mathbf{V}$ est donnée par :

$$dJ(\Gamma, \mathbf{V}) = \int_{\Gamma} (\nabla k(s) \cdot \mathbf{N} - k(s) \kappa) (\mathbf{V} \cdot \mathbf{N}) ds \quad (2.24)$$

avec $\mathbf{N}$ la normale intérieure au contour, κ la courbure moyenne de Γ et s son abscisse curviligne.

A partir de cette dérivée eulérienne, nous pouvons calculer la vitesse d'évolution du contour actif qui fera évoluer la courbe vers un minimum de la fonctionnelle. Comme expliqué à l'équation 2.21, nous choisissons une vitesse de façon à ce que la dérivée du critère soit négative. Nous obtenons la vitesse suivante :

$$\mathbf{v} = (k(\mathbf{s})\kappa - \nabla k(\mathbf{s}) \cdot \mathbf{N})\mathbf{N} \quad \text{avec } \mathbf{s} = \Gamma(p, \tau) \quad (2.25)$$

Ce qui donne pour l'équation d'évolution du contour actif :

$$\begin{cases} \frac{\partial \Gamma(p, \tau)}{\partial \tau} = (k(\mathbf{s})\kappa - \nabla k(\mathbf{s}) \cdot \mathbf{N})\mathbf{N} \\ \Gamma(p, 0) = \Gamma_0(p) \end{cases} \quad (2.26)$$

avec Γ_0 un contour initial, $s = \Gamma(p, \tau)$ et p le paramètre du contour Γ . Cette vitesse a été calculée dans [CKS97].

2.3.3.2 Dérivation d'un terme basé régions dont le descripteur est indépendant de la région

Un terme basé régions dont le descripteur ne dépend pas de la région s'écrit sous la forme :

$$J(\Omega) = \int_{\Omega} k(\mathbf{x}) d\mathbf{x} \quad (2.27)$$

où $k(\mathbf{x})$ est le descripteur de la région.

Nous utilisons le théorème 2.1 pour calculer la dérivée eulérienne de cette fonctionnelle dans la direction $\mathbf{V}$:

$$dJ(\Omega, \mathbf{V}) = \int_{\Omega} k'(\mathbf{x}, \Omega, \mathbf{V}) d\mathbf{x} - \int_{\Gamma} k(s, \Omega) (\mathbf{V} \cdot \mathbf{N}) ds$$

Or ici $k(\mathbf{x}, \Omega) = k(\mathbf{x})$ donc $k'(\mathbf{x}, \Omega, \mathbf{V}) = 0$, alors la dérivée devient :

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} k(s) (\mathbf{V} \cdot \mathbf{N}) ds$$

Nous obtenons alors l'équation d'évolution du contour actif suivante :

$$\begin{cases} \frac{\partial \Gamma(p, \tau)}{\partial \tau} = k(\mathbf{s})\mathbf{N} \\ \Gamma(p, 0) = \Gamma_0(p) \end{cases} \quad (2.28)$$

avec Γ_0 un contour initial, $s = \Gamma(p, \tau)$ et p le paramètre du contour Γ .

2.3.3.3 Dérivation d'un terme basé régions dont le descripteur dépend de la région

Un terme basé régions à double dépendance est l'intégrale sur une région d'un descripteur dépendant lui-même de la région. Il s'écrit :

$$J(\Omega) = \int_{\Omega} k(\mathbf{x}, \Omega) d\mathbf{x} \quad (2.29)$$

avec $k(\mathbf{x}, \Omega)$ le descripteur dépendant de la région Ω .

Nous calculons la dérivée eulérienne du critère (2.29) dans le cas d'un descripteur fonction de la moyenne de l'intensité de la région, pour illustrer la méthode de dérivation. Le descripteur $k(\mathbf{x}, \Omega)$ est de la forme :

$$k(\mathbf{x}, \Omega) = \varphi(I(\mathbf{x}) - \mu(\Omega))$$

avec $\varphi(r)$ une fonction positive de classe $C^1(\mathbb{R})$, paire et croissante sur $\mathbb{R}^+$. La moyenne μ est définie par :

$$\mu(\Omega) = \frac{1}{|\Omega|} \int_{\Omega} I(\mathbf{x}) d\mathbf{x}$$

La fonctionnelle à minimiser s'écrit alors :

$$J(\Omega) = \int_{\Omega} \varphi(I(\mathbf{x}) - \mu(\Omega)) d\mathbf{x} \quad (2.30)$$

Un telle fonctionnelle cherche à segmenter une région qui est homogène puisque l'on cherche à minimiser la distance à la moyenne de la région.

Pour calculer la dérivée du critère (2.30), nous allons utiliser le théorème 2.1. Il sera nécessaire d'appliquer plusieurs fois ce théorème de façon récursive afin d'obtenir uniquement des intégrales de contour. Nous appliquons une première fois ce théorème à la fonctionnelle (2.30) et nous obtenons la formule suivante :

$$dJ(\Omega, \mathbf{V}) = \int_{\Omega} \varphi'(I(\mathbf{x}) - \mu(\Omega), \mathbf{V}) d\mathbf{x} - \int_{\Gamma} \varphi(I(s) - \mu(\Omega)) (\mathbf{V} \cdot \mathbf{N}) ds \quad (2.31)$$

La dérivée de domaine $\varphi'(I(\mathbf{x}) - \mu(\Omega), \mathbf{V})$ peut se décomposer en :

$$\varphi'(I(\mathbf{x}) - \mu(\Omega), \mathbf{V}) = -\mu'(\Omega, \mathbf{V}) \varphi'(I(\mathbf{x}) - \mu(\Omega))$$

avec $\varphi'(r)$ la dérivée de φ par rapport à r . Il nous faut donc calculer la dérivée eulérienne de $\mu(\Omega)$ dans la direction $\mathbf{V}$.

Nous pouvons écrire la fonction μ de la façon suivante :

$$\mu(\Omega) = \frac{1}{|\Omega|} \int_{\Omega} I(\mathbf{x}) d\mathbf{x} = f(G_1(\Omega), G_2(\Omega)) \quad (2.32)$$

avec

$$\begin{aligned} G_1(\Omega) &= \int_{\Omega} I(\mathbf{x}) d\mathbf{x} \\ G_2(\Omega) &= \int_{\Omega} d\mathbf{x} = |\Omega| \\ f(G_1(\Omega), G_2(\Omega)) &= \frac{G_1(\Omega)}{G_2(\Omega)} \end{aligned}$$

Calculons la dérivée eulérienne de la fonction f sans la direction $\mathbf{V}$:

$$df(\Omega, \mathbf{V}) = dG_1(\Omega, \mathbf{V}) \frac{\partial f}{\partial G_1}(G_1, G_2) + dG_2(\Omega, \mathbf{V}) \frac{\partial f}{\partial G_2}(G_1, G_2) \quad (2.33)$$

Les dérivées partielles de f par rapport à G_1 et G_2 sont :

$$\frac{\partial f}{\partial G_1} = \frac{1}{G_2} = \frac{1}{|\Omega|} \quad (2.34)$$

$$\frac{\partial f}{\partial G_2} = -\frac{G_1}{G_2^2} = -\frac{\int_{\Omega} I(\mathbf{x}) d\mathbf{x}}{|\Omega|^2} = -\frac{\mu(\Omega)}{|\Omega|} \quad (2.35)$$

Nous devons maintenant calculer les dérivées eulériennes de G_1 et G_2 dans la direction $\mathbf{V}$. Nous appliquons à nouveau le théorème 2.1. La dérivée eulérienne de G_1 dans la direction $\mathbf{V}$ vaut :

$$dG_1(\Omega, \mathbf{V}) = \int_{\Omega} I'(\mathbf{x}, \Omega, \mathbf{V}) d\mathbf{x} - \int_{\Gamma} I(s)(\mathbf{V} \cdot \mathbf{N}) ds \quad (2.36)$$

$I(\mathbf{x})$ est constante par rapport à τ donc $I'(\mathbf{x}, \Omega, \mathbf{V}) = 0$. Donc la dérivée eulérienne de G_1 est :

$$dG_1(\Omega, \mathbf{V}) = - \int_{\Gamma} I(s)(\mathbf{V} \cdot \mathbf{N}) ds \quad (2.37)$$

De même, nous calculons la dérivée eulérienne de G_2 dans la direction $\mathbf{V}$:

$$dG_2(\Omega, \mathbf{V}) = \int_{\Omega} 1'(\mathbf{x}, \Omega, \mathbf{V}) d\mathbf{x} - \int_{\Gamma} (\mathbf{V} \cdot \mathbf{N}) ds \quad (2.38)$$

avec $1(\mathbf{x})$ la fonction constante 1 donc sa dérivée est nulle. La dérivée de G_2 se réduit donc à :

$$dG_2(\Omega, \mathbf{V}) = - \int_{\Gamma} (\mathbf{V} \cdot \mathbf{N}) ds \quad (2.39)$$

Ensuite en regroupant (2.34), (2.35), (2.37) et (2.38), nous avons la dérivée eulérienne de $\mu(\Omega)$ dans la direction $\mathbf{V}$:

$$d\mu(\Omega, \mathbf{V}) = -\frac{1}{|\Omega|} \int_{\Gamma} I(s)(\mathbf{V} \cdot \mathbf{N}) ds + \frac{\mu(\Omega)}{|\Omega|} \int_{\Gamma} (\mathbf{V} \cdot \mathbf{N}) ds \quad (2.40)$$

$$= - \int_{\Gamma} \frac{I(s) - \mu(\Omega)}{|\Omega|} (\mathbf{V} \cdot \mathbf{N}) ds \quad (2.41)$$

Finalement en injectant la dérivée de domaine de φ et la dérivée eulérienne de μ dans l'équation (2.31), nous obtenons :

$$dJ(\Omega, \mathbf{V}) = \int_{\Omega} \left(\int_{\Gamma} \frac{I(s) - \mu(\Omega)}{|\Omega|} (\mathbf{V} \cdot \mathbf{N}) ds \right) \varphi'(I(\mathbf{x}) - \mu(\Omega)) d\mathbf{x} \quad (2.42)$$

$$- \int_{\Gamma} \varphi(I(s) - \mu(\Omega)) (\mathbf{V} \cdot \mathbf{N}) ds \quad (2.43)$$

Ce qui donne en inversant les intégrales :

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} \left(\varphi(I(s) - \mu(\Omega)) - \frac{I(s) - \mu(\Omega)}{|\Omega|} \int_{\Omega} \varphi'(I(\mathbf{x}) - \mu(\Omega)) d\mathbf{x} \right) (\mathbf{V} \cdot \mathbf{N}) ds \quad (2.44)$$

L'équation d'évolution déduite de cette dérivée est alors :

$$\begin{cases} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= \left(\varphi(I(\mathbf{s}) - \mu(\Omega)) - \frac{I(\mathbf{s}) - \mu(\Omega)}{|\Omega|} \int_{\Omega} \varphi'(I(\mathbf{x}) - \mu(\Omega)) d\mathbf{x} \right) \mathbf{N} \\ \Gamma(p, 0) &= \Gamma_0(p) \end{cases} \quad (2.45)$$

avec $\mathbf{s} = \Gamma(p, \tau)$.

Dans le cas particulier où $\varphi(r) = r^2$, le critère s'écrit :

$$J(\Omega) = \int_{\Omega} (I(\mathbf{x}) - \mu(\Omega))^2 d\mathbf{x} \quad (2.46)$$

Et alors la dérivée eulérienne vaut :

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} \left((I(s) - \mu(\Omega))^2 - \frac{I(s) - \mu(\Omega)}{|\Omega|} \int_{\Omega} 2(I(\mathbf{x}) - \mu(\Omega)) d\mathbf{x} \right) (\mathbf{V} \cdot \mathbf{N}) ds \quad (2.47)$$

Cependant :

$$\int_{\Omega} 2(I(\mathbf{x}) - \mu(\Omega)) d\mathbf{x} = 0$$

Donc la dérivée se réduit à :

$$dJ(\Omega, \mathbf{V}) = - \int_{\Gamma} \left((I(s) - \mu(\Omega))^2 (\mathbf{V} \cdot \mathbf{N}) \right) ds \quad (2.48)$$

Ce qui donne pour l'équation d'évolution du contour actif :

$$\begin{cases} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= (I(\mathbf{s}) - \mu(\Omega))^2 \mathbf{N} \\ \Gamma(p, 0) &= \Gamma_0(p) \end{cases} \quad (2.49)$$

avec $\mathbf{s} = \Gamma(p, \tau)$.

Nous pourrions développer ce principe de dérivation pour d'autres descripteurs, par exemple un descripteur fonction de la variance. Les principes de dérivation restent les

mêmes, à savoir appliquer de façon récursive le théorème de dérivation 2.1 afin de retrouver de façon équivalente une dérivée ne contenant plus que des termes d'intégrales de contour. Dans l'exemple présenté nous avons vu qu'il suffisait d'appliquer deux fois le théorème avant de parvenir à l'écriture de la dérivée eulérienne de façon satisfaisante.

Le lecteur pourra trouver une version plus générale du théorème 2.1 dans la thèse de S. Jehan-Besson [JB03] que nous ne détaillerons pas ici.

Une fois obtenue l'équation d'évolution du contour actif, voyons comment mettre en œuvre l'algorithme des contours actifs.

2.4 IMPLÉMENTATION

L'implémentation des équations aux dérivées partielles qui régissent les contours actifs n'est pas simple à mettre en place. Elle a besoin d'être stable et précise. Il faut aussi que l'approximation lagrangienne de l'équation d'évolution ait un pas très petit. De plus, il faut pouvoir gérer les changements de topologie : un contour doit pouvoir se séparer en plusieurs contours tout comme deux contours doivent pouvoir fusionner.

Deux approches principales existent pour la mise en œuvre des contours actifs : les approches explicites et les approches implicites. Une approche explicite définit le contour directement à partir de paramètres tandis qu'une approche implicite comme les ensembles de niveaux (*level sets*) définit le contour de manière implicite comme étant le niveau zéro d'une fonction de dimension supérieure.

Nous présentons rapidement la méthode des ensembles des niveaux et l'approche paramétrique utilisant des B-splines pour comparer les avantages et inconvénients de chaque approche.

2.4.1 Approche implicite : les ensembles de niveaux

La méthode des ensembles de niveaux est une approche dite implicite d'implémentation des contours actifs. Elle a été popularisée par Osher et Sethian [OS88]. Son principe consiste à considérer que le contour actif est le niveau zéro d'une fonction de dimension supérieure $u : \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R}$. Nous avons donc :

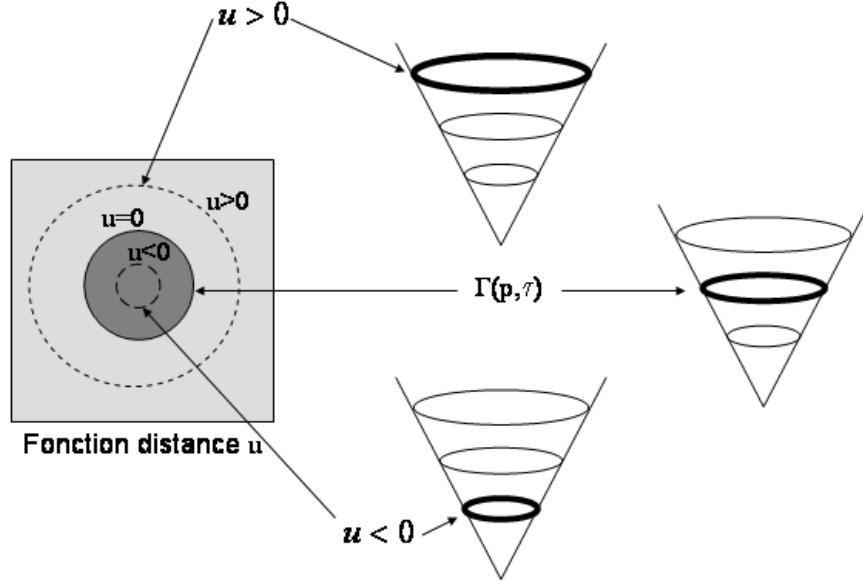
$$\mathbf{x} \in \Gamma(p, \tau) \Leftrightarrow u(\mathbf{x}, \tau) = 0$$

Ce qui peut s'écrire aussi :

$$u(\Gamma(p, \tau), \tau) = 0, \quad \forall p \in [a, b], \forall \tau \geq 0 \quad (2.50)$$

Nous pouvons choisir u comme étant une fonction distance signée au contour $\Gamma(p, \tau)$, négative à l'intérieure de la courbe et positive à l'extérieur (voir la figure 2.2). Nous avons alors presque partout : $|\nabla u| = 1$.

Les avantages d'une telle méthode sont les suivants [Set96] :

FIG. 2.2 – Illustration de la fonction distance signée u

- Les changements de topologie sont automatiquement gérés.
- La fonction distance u permet des schémas numériques stables et précis.
- Les propriétés géométriques du contour peuvent être facilement estimées à partir de la fonction u .
- La formulation peut être aisément étendue aux dimensions supérieures.

Rappelons que l'évolution du contour actif est la suivante :

$$\frac{\partial \Gamma(p, \tau)}{\partial \tau} = F(p, \tau) \mathbf{N}(p, \tau)$$

avec F l'amplitude de la vitesse d'évolution du contour actif et $\mathbf{N}$ la direction normale au contour. Nous pouvons dériver u (équation (2.50)) par rapport à τ , ce qui nous donne :

$$\frac{\partial u(\Gamma(p, \tau))}{\partial \tau} = F(p, \tau) |\nabla u(\Gamma(p, \tau), \tau)| \quad (2.51)$$

Cette équation est valide uniquement pour le niveau zéro mais nous pouvons l'étendre sur l'image si F y est définie. Nous obtenons donc :

$$\frac{\partial u(\mathbf{x}, \tau)}{\partial \tau} = F(\mathbf{x}, \tau) |\nabla u(\mathbf{x}, \tau)| \quad (2.52)$$

avec $u(\mathbf{x}, 0)$ la fonction distance de Γ_0 .

Le calcul de u dans (2.52) pose des problèmes d'instabilités, en particulier parce que u ne conserve pas les propriétés de distance, et a donc un gradient qui peut tendre vers

l'infini. Une solution à ce problème peut être de réinitialiser la fonction distance avec une EDP de telle façon que $|\nabla u| = 1$ au cours de l'évolution :

$$\frac{\partial u(\mathbf{x}, \tau)}{\partial \tau} = \text{sign}(u(\mathbf{x}, \tau))(1 - |\nabla u(\mathbf{x}, \tau)|) \quad (2.53)$$

avec $\text{sign}(u(\mathbf{x}, \tau))$ une fonction caractéristique du signe de $u(\mathbf{x}, \tau)$.

Cette méthode a un coût calcul important, c'est pourquoi a été introduite par Chopp [Cho91] la méthode dite de la bande étroite (ou *narrowband* en anglais). Cela consiste à ne faire les calculs que sur une bande entourant le niveau zéro de u et à remettre à jour cette bande à chaque fois que la courbe s'approche de ses bords. Ainsi, les coûts calcul sont nettement réduits. Cette méthode a été utilisée notamment par [MSV95, AS95, Set96].

Cependant, cette approche reste tout de même très lourde en calculs, et ne permet pas de convertir facilement le contour en courbe géométrique orientée. D'autres approches ont été développées, notamment les modèles paramétriques de contours actifs, permettant de limiter les temps de calcul. Nous présentons maintenant une approche explicite d'implémentation des contours actifs, en considérant un modèle paramétrique pour ces contours.

2.4.2 Approche explicite : les modèles paramétriques

Les méthodes paramétriques définissent le contour actif à partir d'un ensemble de paramètres. Parmi les premiers modèles paramétriques, nous pouvons citer le modèle des *snakes* de Kass *et al.* [KWT88]. Les propriétés intrinsèques du modèle sont choisies afin de minimiser au mieux la fonctionnelle. Les contours actifs sont définis comme des courbes polygonales soumises à des forces issues d'énergies basées notamment sur les propriétés du contour. Cette approche a l'inconvénient de nécessiter une courbe initiale proche de la solution.

De nombreux travaux ont été développés afin d'améliorer cette représentation des contours actifs. Menet *et al.* [MSMM90] ont étendu le modèle polygonal des *snakes* en proposant une structure utilisant les fonctions B-spline. Ces fonctions permettent d'obtenir des courbes avec une régularité qui permet de calculer directement les grandeurs géométriques du contour actif. De plus, ces courbes apportent une flexibilité, une simplicité de modélisation et de bonnes propriétés de régularité qui ont entraîné un grand intérêt pour ce type de modélisation. C'est dans ce cadre que Cohen *et al.* [Coh91, CBA93] ont introduit la force ballon, applicable à ce genre de courbes. Récemment, Unser *et al.* [Uns99, JBU01] ont étudié de façon approfondie les fonctions splines et sont parvenus à élaborer un modèle de contours actifs exploitant les propriétés intrinsèques de ces courbes. Puis Precioso *et al.* [PBBU05] ont utilisé des B-splines cubiques pour représenter les contours actifs. C'est cette dernière approche que nous utilisons dans ce manuscrit.

Les B-splines cubiques sont des fonctions deux fois continûment différentiables, formées de polygones de degré trois. Elles sont définies par des points de contrôle Q_i et ont la propriété intéressante d'être locales : déplacer un point de contrôle ne modifie que quatre intervalles de la courbe. Notons $S_i(t)$ la courbe paramétrée par t entre les point P_i et P_{i+1} ,

la courbe Γ passant par une suite de n points P_i . L'utilisation de B-splines cubique permet d'assurer une continuité C^2 même aux points P_i . Chaque segment de la courbe est exprimé selon des polynômes de degré trois :

$$S_i(t) = Q_i B_i^3(t) + Q_{i+1} B_{i+1}^3(t) + Q_{i+2} B_{i+2}^3(t) + Q_{i+3} B_{i+3}^3(t)$$

où les $B_i^3(t)$ sont des fonctions cubiques B-splines non uniformes, c'est-à-dire des fonctions polynômiales de degré trois, définies par morceaux. Les points de contrôle Q_i sont estimés à partir des points d'interpolation P_i de la spline. La construction de la spline est présentée sur la figure 2.3.

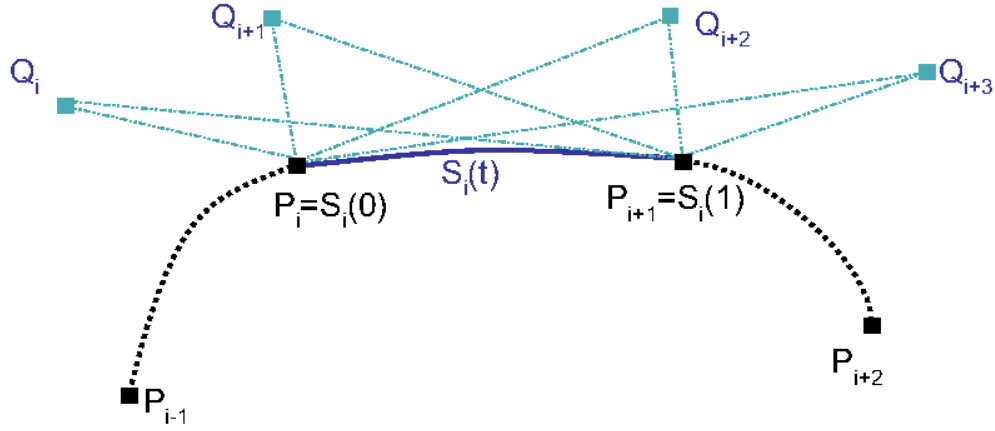


FIG. 2.3 – Construction d'une B-spline cubique

Reprenons l'évolution du contour actif :

$$\frac{\partial \Gamma(p, \tau)}{\partial \tau} = F(p, \tau) \mathbf{N}(p, \tau)$$

La force F est appliquée à chaque point d'interpolation P_i , dans la direction de la normale. Une nouvelle spline est alors définie et ainsi de suite jusqu'à convergence de l'algorithme.

Ces approches paramétriques ont l'avantage principal de réduire considérablement le temps d'exécution. Cependant les changements de topologie ne sont plus gérés automatiquement comme cela était le cas avec la méthode des ensembles de niveaux et vont devoir être pris en compte dans l'implémentation.

Dans ce manuscrit, nous présentons des résultats obtenus avec une implémentation paramétrique utilisant des B-splines, principalement pour le coût calcul. De ce fait, nous

n'avons pas besoin de rajouter à notre critère de segmentation une pénalité de régularisation, comme par exemple une contrainte sur la longueur du contour, puisque la régularité de la courbe est réglée par le nombre de points d'interpolation de la courbe.

2.5 CONCLUSION

Les contours actifs basés régions apparaissent comme de puissants outils pour la segmentation d'images et de vidéos. Les méthodes d'obtention de l'équation d'évolution du contour actif sont variées mais nous nous intéressons essentiellement à des méthodes variationnelles, qui consistent en la minimisation d'un critère. Nous utilisons la notion de gradient de forme pour dériver ce critère pouvant comporter des termes basés régions et des termes basés contour et nous obtenons ainsi l'équation d'évolution du contour actif.

Les critères de segmentation sont souvent des fonctions de caractéristiques de l'image (comme son intensité) qui supposent une distribution prédéfinie de ces caractéristiques. Dans ce document, nous nous délivrons de cette contrainte et nous considérons directement les "vraies" distributions des données dans la mesure où nous ne faisons aucune hypothèse sur ces distributions. Notre objectif est d'étudier des critères issus de la théorie de l'information dans un contexte non-paramétrique, c'est-à-dire ne faisant aucune hypothèse sur les distributions des données, pour la segmentation d'images et de vidéos par contours actifs.

Critères statistiques provenant de la théorie de l'information

De nombreuses méthodes de contours actifs présentées dans le chapitre 2 utilisent des critères de segmentation qui font une hypothèse implicite sur les distributions des données. Par exemple, utiliser un critère reposant sur la moyenne de l'intensité d'une région revient à faire une hypothèse gaussienne sur l'intensité de la région. Dans de nombreux cas réels, ces hypothèses ne sont pas vérifiées et introduisent donc des erreurs. Dans nos travaux, nous ne faisons aucune hypothèse sur les distributions des données que nous prenons en compte et de ce fait nous autorisons des distributions très générales.

Nous proposons de minimiser des fonctions de ces distributions non-paramétriques en s'inspirant des mesures provenant de la théorie de l'information. Ainsi nous ne faisons plus d'hypothèse sur les distributions des données que nous considérons et nous estimons ces distributions avec des méthodes non-paramétriques qui nous permettent de nous approcher au mieux des "vraies" distributions.

3.1 INTRODUCTION

La théorie de l'information intervient notamment dans les domaines suivants : théorie de la communication, physique, statistiques, mathématiques, économie, informatique, théorie des probabilités. Depuis Shannon [Sha48], elle a montré son utilité dans bon nombre d'applications, de la transmission de données aux calculs d'investissement, de la thermodynamique au calcul de complexité d'une chaîne de données, etc ...

Une notion de base de la théorie de l'information est l'entropie. C'est une mesure d'information. En compression de données, l'entropie d'une variable aléatoire représente la borne minimale de la longueur moyenne des mots-code pour pouvoir décrire complètement la variable aléatoire. Cette notion est à mettre en parallèle avec la notion d'entropie en thermodynamique où selon la deuxième loi de la thermodynamique, l'entropie d'un système isolé ne peut qu'augmenter ou rester constante. De manière générale, l'entropie d'un signal correspond à la quantité d'information contenue dans le signal. Plus le signal est redondant en terme de probabilités, plus son entropie est faible. Plus les probabilités de chaque valeur du signal sont égales, plus l'entropie est forte.

D'autres mesures d'information sont à mettre en relation avec l'entropie. Il s'agit tout d'abord de l'information mutuelle qui mesure la dépendance entre deux variables aléatoires. L'information mutuelle est un cas particulier d'une notion appelée entropie relative qui mesure une "distance" entre deux fonctions de distributions. Bien que cela ne soit pas une vraie distance, l'entropie relative possède quelques propriétés des distances, par exemple elle est toujours positive et égale à zéro seulement si les deux fonctions de distributions sont égales.

Dans ce chapitre nous verrons comment nous pouvons utiliser ces notions de théorie de l'information dans un schéma de contours actifs et quelles mesures sont les plus pertinentes pour effectuer une segmentation d'images. Nous introduirons tout d'abord un critère général prenant en compte des caractéristiques statistiques de l'image puis nous le déclinerons de façon à utiliser différentes mesures d'information. Nous présenterons tout d'abord un critère basé sur l'entropie, puis l'entropie jointe et l'information mutuelle. Enfin nous introduirons la divergence de Kullback-Leibler.

3.2 CRITÈRE GÉNÉRAL BASÉ SUR LES DISTRIBUTIONS

Nous voulons définir un critère général de segmentation, prenant en considération les caractéristiques statistiques de l'image. Notons $\mathbf{f}$ une caractéristique de l'image. Cette caractéristique peut être par exemple l'intensité lumineuse d'une image en niveaux de gris ou bien les canaux couleurs d'une image couleur. L'idée de cette partie est de ne pas prendre un modèle défini *a priori* pour la distribution de cette caractéristique comme cela est fréquemment le cas [CV01, YTW99, JBBA03]. Une hypothèse commune est de considérer que les caractéristiques de l'image suivent une loi connue, par exemple une loi gaussienne, laplacienne ou de Rayleigh. Pour notre part, ne connaissant pas les règles régissant la génération des images réelles, nous ne voulons pas imposer de loi prédéfinie pour la distribution des caractéristiques considérées (notamment l'intensité). D'autre part, ces densités de probabilité sont en général unimodales alors que les densités de probabilité réellement rencontrées s'avèrent être plus souvent multi-modales. Nous ferons par contre l'hypothèse que les pixels voisins ont des caractéristiques (notamment l'intensité lumineuse) indépendantes. Nous supposerons aussi que les différents canaux des densités de probabilité de ces caractéristiques sont indépendants.

3.2.1 Estimation d'une densité de probabilité

L'estimation de paramètres est un domaine à part entière dans le traitement du signal. De nombreuses études ont permis d'avoir des méthodes efficaces et sans biais. Le spectre de ces techniques est assez large. Nous nous intéressons aux méthodes non-paramétriques pour l'estimation de densités de probabilités. Un outil communément répandu dans les domaines du signal est l'estimateur à noyau, introduit par Rosenblatt [Ros56] et Parzen [Par62]. De nombreux ouvrages abordent ce type de construction [Sco92, Sil86, DL01, DHS00]. D'autres méthodes comme celles des k -ièmes plus proches voisins [DT77, Sil86] peuvent aussi être utilisées mais elles ne feront pas l'objet d'étude dans ce manuscrit car nous nous intéressons aux densités continues lisses dont nous pouvons avoir une expression analytique. La méthode de Parzen utilise une densité-noyau K et un paramètre h , le paramètre d'échelle

de la densité K , qui détermine le niveau de lissage de l'estimation par la largeur des noyaux. Ce noyau K doit avoir les propriétés suivantes :

- K doit être symétrique, $K(-u) = K(u)$
- $\int_{\mathbb{R}} K(u) du = 1$
- $\int_{\mathbb{R}} u^j K(u) du = 0$ pour $j = 1, \dots, k-1$
- $\int_{\mathbb{R}} u^k K(u) du \neq 0$

K est alors appelé noyau d'ordre k . Fréquemment on choisit un noyau d'ordre 2, et notamment le noyau gaussien de moyenne nulle (noyau centré). La formule d'un tel estimateur est donnée par :

$$\hat{p}_h(x) = \frac{1}{N} \sum_{i=1}^N K_h(x - x_i) \quad (3.1)$$

où $\hat{p}_h$ est l'approximation de la densité de probabilité, h est la largeur de bande, $K_h(u) = K(u/h)/h$, N le nombre de points de l'échantillon et en choisissant un noyau gaussien centré, K est donné par :

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) \quad (3.2)$$

La figure 3.1 illustre le principe de construction d'un estimateur à noyaux : plus il y a de noyaux dans un voisinage, et plus la densité de points y est élevée.

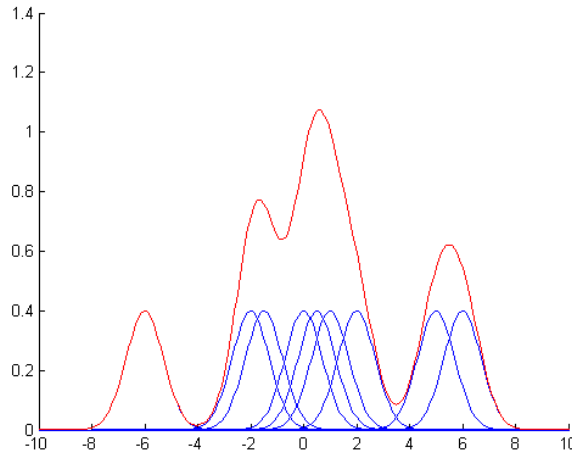


FIG. 3.1 – Illustration de la méthode de Parzen. Gaussiennes avec $h = 1$ en bleu et somme de ces gaussiennes en rouge. L'estimation de la densité de probabilité de Parzen est donnée en divisant cette somme par le nombre de gaussiennes.

Le choix du paramètre h est crucial dans l'estimation de la densité de probabilité résultante. S'il est trop grand, la densité estimée sera trop lisse, s'il est trop petit la densité estimée sera trop perturbée. De nombreuses études se sont intéressées à la façon de déterminer ce paramètre [BD94, Tür94, Dev94] mais sa détermination reste difficile, notamment

pour des dimensions supérieures à 1. Il peut être calculé de façon adaptée en utilisant la déviation standard des données [Sil86]. Il existe aussi des méthodes utilisant une taille de noyau variable (pas la même largeur de bande pour tous les points de l'échantillon) mais leur coût calcul est assez important [Sil86, Sco92, BD94, DL01].

Les figures 3.2, 3.3 et 3.4 montrent une estimation de la densité de probabilité (que l'on notera parfois *pdf* : *probability density function*, en anglais) d'un échantillon de 100 observations de densité $\frac{1}{2}\mathcal{N}(-4, 1) + \frac{1}{2}\mathcal{N}(0, 2)$ avec une largeur de bande de 0.1 (figure 3.2) et de 1 (figure 3.3). Avec un h trop petit, la *pdf* est mal estimée et trop perturbée ; avec un h trop grand la *pdf* est trop lissée et aplatie. Sur la figure 3.4, la largeur de bande est calculée automatiquement en fonction des données et donne une densité de probabilités proche des données réelles (méthode *plug-in*).

Ces figures montrent bien l'influence de la largeur de bande sur la densité de probabilité estimée et le soin qu'il faut apporter à son évaluation.

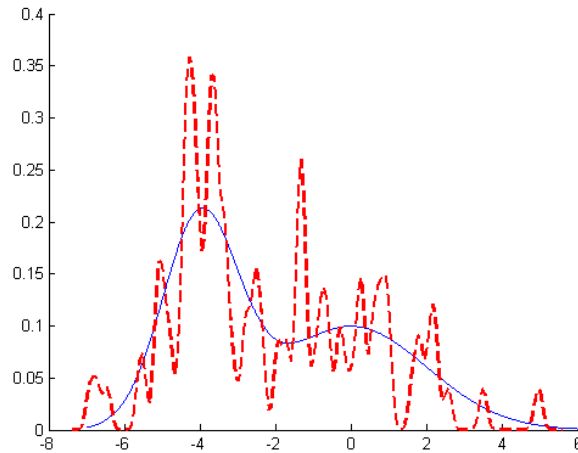


FIG. 3.2 – Illustration de la méthode de Parzen, influence de la largeur de bande h . La ligne bleue continue représente la densité à estimer, la ligne rouge en pointillés la densité obtenue avec une largeur de bande de 0.1.

L'estimation de densité de probabilités par méthode à noyau est très utilisée pour des données de dimension 1 mais ce type d'estimateur devient rapidement inefficace quand la dimension des données augmente [Sco92] (pour une dimension supérieure à 2 ou 3). De plus, des techniques comme les k -ièmes plus proches voisins semblent plus adaptées pour ces cas.

3.2.2 Critère fonction de la densité de probabilité

Nous allons donc considérer pour la segmentation un critère qui prend en compte ces densités de probabilité afin de coller au mieux aux vraies caractéristiques de l'image. No-

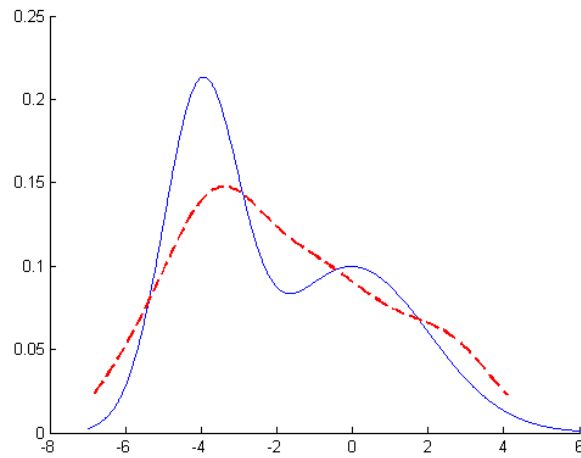


FIG. 3.3 – Illustration de la méthode de Parzen, influence de la largeur de bande h . La ligne bleue continue représente la densité à estimer, la ligne rouge en pointillés la densité obtenue avec une largeur de bande de 1.

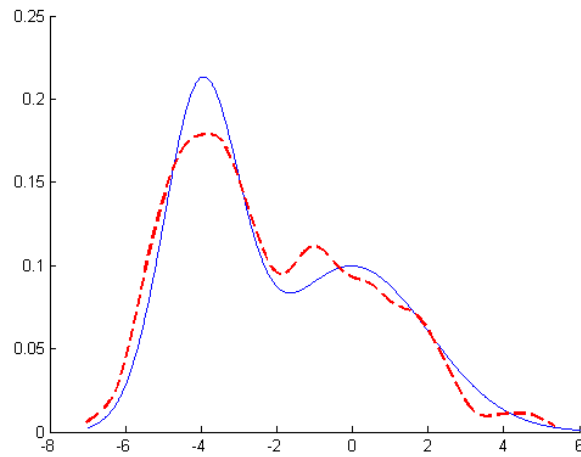


FIG. 3.4 – Illustration de la méthode de Parzen, influence de la largeur de bande h . La ligne bleue continue représente la densité à estimer, la ligne rouge en pointillés la densité obtenue avec une largeur de bande calculée à partir des données.

tons $\mathbf{f}$ la caractéristique de l'image considérée. Cette caractéristique peut être l'intensité lumineuse de l'image, un vecteur mouvement ou bien un descripteur de forme, et est de la forme : $\mathbf{f} : \Omega_f \rightarrow \mathbb{R}^m$ où $\Omega_f \subset \mathbb{R}^2$, m est la dimension de la caractéristique $\mathbf{f}$. Si $\mathbf{f}$ est l'intensité de l'image, m représente la dimension de son espace couleur : pour une image en niveaux de gris $m = 1$, pour une image couleur $m = 3$.

Considérons une portion Ω de l'image. La densité de probabilité de la caractéristique de cette région est notée $p(\mathbf{f}(\mathbf{x}), \Omega)$ où $\mathbf{x}$ est un pixel de la région Ω .

Nous choisissons le descripteur de la région suivant :

$$k(\mathbf{x}, \Omega) = \varphi(p(\mathbf{f}(\mathbf{x}), \Omega)) \quad (3.3)$$

où φ est une fonction : $\mathbb{R}^+ \rightarrow \mathbb{R}^+$ de la *pdf* de la caractéristique dans la région Ω .

Nous considérons la région Ω et estimons la *pdf* en utilisant la méthode des noyaux de Parzen, en continu :

$$p(\mathbf{f}(\mathbf{x}), \Omega) = \frac{1}{|\Omega|h} \int_{\Omega} K\left(\frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\hat{\mathbf{x}})}{h}\right) d\hat{\mathbf{x}} \quad (3.4)$$

où K est le noyau gaussien de dimension m , h est sa largeur de bande et $|\Omega|$ est la mesure de la région Ω ; pour $m = 1$:

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right).$$

La largeur de bande peut être exprimée en fonction de la mesure de la région afin de garantir la convergence en probabilités vers la vraie *pdf*. Prenons par exemple $h = h_1/\sqrt{|\Omega|}$, avec h_1 une constante de normalisation. Cela signifie que plus la mesure de la région Ω est grande, plus le noyau est étroit, et inversement. :

$$\begin{aligned} p(\mathbf{f}(\mathbf{x}), \Omega) &= \frac{1}{|\Omega|\frac{h_1}{\sqrt{|\Omega|}}} \int_{\Omega} K\left(\frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\hat{\mathbf{x}})}{\frac{h_1}{\sqrt{|\Omega|}}}\right) d\hat{\mathbf{x}} \\ &= \frac{1}{\sqrt{|\Omega|}h_1} \int_{\Omega} K\left(\frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\hat{\mathbf{x}})}{\frac{h_1}{\sqrt{|\Omega|}}}\right) d\hat{\mathbf{x}} \end{aligned} \quad (3.5)$$

Dans nos travaux, nous ne tiendrons pas compte de la dépendance de la largeur de bande dans la dérivation du critère car le nombre d'échantillons, c'est-à-dire d'éléments de la région Ω varie relativement peu au cours du processus de segmentation. Nous reprenons donc la définition donnée en (3.4).

Le critère général de segmentation est donné par :

$$J(\Omega) = \int_{\Omega} k(\mathbf{x}, \Omega) d\mathbf{x} = \int_{\Omega} \varphi(p(\mathbf{f}(\mathbf{x}), \Omega)) d\mathbf{x} \quad (3.6)$$

avec $p(\mathbf{f}(\mathbf{x}))$ la *pdf* de la caractéristique $\mathbf{f}$ au pixel $\mathbf{x}$.

3.2.3 Dérivation du critère fonction de la pdf

Nous avons besoin de dériver le critère (3.6) pour obtenir l'équation d'évolution du contour actif. En utilisant les outils de dérivation décrits dans la section 2.3, nous obtenons la dérivée eulérienne dJ du critère (3.6) dans la direction $\mathbf{V}$:

$$dJ(\Omega, \mathbf{V}) = \int_{\Omega} \varphi'(p(\mathbf{f}(\mathbf{x}), \Omega), \mathbf{V}) d\mathbf{x} - \int_{\Gamma} \varphi(p(\mathbf{f}(s), \Omega)) (\mathbf{V} \cdot \mathbf{N}) ds \quad (3.7)$$

où $\varphi'(p(\mathbf{f}(\mathbf{x}), \Omega), \mathbf{V})$ est la dérivée de domaine de φ dans la direction $\mathbf{V}$ et $\mathbf{N}$ est le vecteur normal unitaire dirigé vers l'intérieur du contour. Le premier terme de l'intégrale représente la dépendance du descripteur φ à la région Ω tandis que le second terme vient de l'évolution de la région Ω elle-même.

Pour déterminer la dérivée de domaine de φ , nous avons besoin de dériver la pdf p . Pour rappel l'expression de p est donnée par :

$$p(\mathbf{f}(\mathbf{x}), \Omega) = \frac{1}{|\Omega|} \int_{\Omega} K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(\hat{\mathbf{x}})) d\hat{\mathbf{x}}. \quad (3.8)$$

Nous pouvons dériver p comme un produit. Notons $\mathcal{A} = \frac{1}{|\Omega|}$ et $\mathcal{B} = \int_{\Omega} K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(\hat{\mathbf{x}})) d\hat{\mathbf{x}}$. Nous avons :

$$\mathcal{A}(\Omega) = \frac{1}{|\Omega|} = \frac{1}{\int_{\Omega} d\mathbf{x}} = \frac{G_1(\Omega)}{G_2(\Omega)} = f(G_1(\Omega), G_2(\Omega)). \quad (3.9)$$

La dérivée eulérienne de $\mathcal{A}$ est donnée par :

$$d\mathcal{A}(\Omega, \mathbf{V}) = dG_1(\Omega, \mathbf{V}) \frac{\partial f}{\partial G_1}(G_1, G_2) + dG_2(\Omega, \mathbf{V}) \frac{\partial f}{\partial G_2}(G_1, G_2). \quad (3.10)$$

Les dérivées partielles de f se calculent facilement :

$$\frac{\partial f}{\partial G_1} = \frac{1}{G_2} = \frac{1}{|\Omega|} \quad (3.11)$$

$$\frac{\partial f}{\partial G_2} = -\frac{G_1}{G_2^2} = -\frac{1}{|\Omega|^2}. \quad (3.12)$$

Pour calculer les dérivées eulériennes de G_1 et G_2 dans la direction $\mathbf{V}$, nous utilisons à nouveau le théorème des gradients de forme :

$$dG_1(\Omega, \mathbf{V}) = 0 \quad (3.13)$$

et

$$dG_2(\Omega, \mathbf{V}) = \int_{\Omega} \mathbf{1}'(\mathbf{x}, \Omega, \mathbf{V}) d\mathbf{x} - \int_{\Gamma} (\mathbf{V} \cdot \mathbf{N}) ds = - \int_{\Gamma} (\mathbf{V} \cdot \mathbf{N}) ds \quad (3.14)$$

où $\mathbf{1}(\mathbf{x})$ est la fonction constante égale à 1.

A partir des équations (3.11), (3.12), (3.13) et (3.14), nous pouvons déterminer la dérivée eulérienne de $\mathcal{A}$ dans la direction $\mathbf{V}$:

$$d\mathcal{A}(\Omega, \mathbf{V}) = \frac{\int_{\Gamma} (\mathbf{V} \cdot \mathbf{N})}{|\Omega|^2} ds. \quad (3.15)$$

Pour dériver $\mathcal{B}$, nous utilisons le théorème des gradients de forme, car ce terme s'écrit sous la forme $\int_{\Omega} k(\mathbf{x}, \Omega) d\mathbf{x}$:

$$\mathcal{B}'(\mathbf{x}, \Omega, \mathbf{V}) = - \int_{\Gamma} K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(s)) ds . \quad (3.16)$$

Nous obtenons donc la dérivée de p :

$$\begin{aligned} p'(\mathbf{f}(\mathbf{x}), \Omega, \mathbf{V}) &= \frac{1}{|\Omega|^2} \int_{\Gamma} (\mathbf{V} \cdot \mathbf{N}) ds \int_{\Omega} K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(s)) ds \\ &\quad - \frac{1}{|\Omega|} \int_{\Gamma} K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(s)) (\mathbf{V} \cdot \mathbf{N}) ds \\ &= \frac{1}{|\Omega|} \int_{\Gamma} p(\mathbf{f}(\mathbf{x}), \Omega) (\mathbf{V} \cdot \mathbf{N}) ds \\ &\quad - \frac{1}{|\Omega|} \int_{\Gamma} K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(s)) (\mathbf{V} \cdot \mathbf{N}) ds . \end{aligned}$$

Finalement,

$$p'(\mathbf{f}(\mathbf{x}), \Omega, \mathbf{V}) = \frac{1}{|\Omega|} \int_{\Gamma} [p(\mathbf{f}(\mathbf{x}), \Omega) - K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(s))] (\mathbf{V} \cdot \mathbf{N}) ds . \quad (3.17)$$

Nous pouvons donc en déduire la dérivée eulérienne de $\varphi(p(\mathbf{f}(\mathbf{x}), \Omega))$ dans la direction $\mathbf{V}$:

$$\begin{aligned} \varphi'(p(\mathbf{f}(\mathbf{x}), \Omega), \mathbf{V}) &= p'(\mathbf{f}(\mathbf{x}), \Omega, \mathbf{V}) \partial_1 \varphi(p(\mathbf{f}(\mathbf{x}), \Omega)) \\ &= \frac{1}{|\Omega|} \int_{\Gamma} \partial_1 \varphi(p(\mathbf{f}(\mathbf{x}), \Omega)) [p(\mathbf{f}(\mathbf{x}), \Omega) - K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(s))] (\mathbf{V} \cdot \mathbf{N}) ds \end{aligned} \quad (3.18)$$

où $\partial_1 \varphi$ est la dérivée de φ par rapport à p .

Nous revenons à la dérivée eulérienne $dJ(\Omega, \mathbf{V})$ du critère (3.6) dans la direction $\mathbf{V}$. D'après le théorème des gradients de forme, $dJ(\Omega, \mathbf{V})$ est de la forme :

$$dJ(\Omega, \mathbf{V}) = \int_{\Omega} p'(\mathbf{f}(\mathbf{x}), \Omega, \mathbf{V}) \partial_1 \varphi(p(\mathbf{f}(\mathbf{x}), \Omega)) d\mathbf{x} - \int_{\Gamma} \varphi(p(\mathbf{f}(s), \Omega)) (\mathbf{V} \cdot \mathbf{N}) ds \quad (3.19)$$

En injectant la dérivée de φ (3.18) dans l'expression précédente, nous obtenons pour la dérivée du critère :

$$\begin{aligned} dJ(\Omega, \mathbf{V}) &= \int_{\Omega} \frac{1}{|\Omega|} \left[\int_{\Gamma} [p(\mathbf{f}(\mathbf{x}), \Omega) - K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(s))] (\mathbf{V} \cdot \mathbf{N}) ds \right] \partial_1 \varphi(p(\mathbf{f}(\mathbf{x}), \Omega)) d\mathbf{x} \\ &\quad - \int_{\Gamma} \varphi(p(\mathbf{f}(s), \Omega)) (\mathbf{V} \cdot \mathbf{N}) ds . \end{aligned} \quad (3.20)$$

Cette dérivée peut s'écrire sous la forme :

$$dJ_r(\Omega, \mathbf{V}) = - \int_{\Gamma} (A(s, \Omega) + \varphi(p(\mathbf{f}(s), \Omega))) (\mathbf{V} \cdot \mathbf{N}) ds \quad (3.21)$$

où $A(s, \Omega)$ est un terme venant de la dépendance du descripteur à la région :

$$A(s, \Omega) = -\frac{1}{|\Omega|} \int_{\Omega} \partial_1 \varphi(p(\mathbf{f}(\mathbf{x}), \Omega)) \left[p(\mathbf{f}(\mathbf{x}), \Omega) - K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(s)) \right] d\mathbf{x}$$

De cette dérivée eulérienne nous pouvons déduire la vitesse d'évolution $\mathbf{v}$ du contour actif qui fera décroître le critère vers un minimum. Nous obtenons l'équation d'évolution suivante :

$$\frac{\partial \Gamma}{\partial \tau}(\mathbf{s}) = \mathbf{v} = \left(A(\mathbf{s}, \Omega) + \varphi(q(\mathbf{f}(\mathbf{s}), \Omega)) \right) \mathbf{N} \quad (3.22)$$

où $\mathbf{s} = \Gamma(p, \tau)$.

Nous avons donc présenté un critère général de segmentation basé sur la densité de probabilités d'une caractéristique de l'image. Pour s'approcher au mieux des vraies distributions, nous avons estimé cette densité par une méthode non-paramétrique à noyau. En utilisant les gradients de forme, nous avons dériver ce critère général afin de parvenir à l'équation d'évolution du contour actif. Etudions maintenant quels critères peuvent être utilisés dans ce cadre général.

3.3 ENTROPIE

Dans cette section, nous nous proposons d'étudier une mesure d'information particulière, à savoir l'entropie. Nous expliquons pourquoi utiliser l'entropie et nous présentons la façon de l'estimer. Nous introduisons ensuite les notions d'entropie conditionnelle et d'entropie jointe.

3.3.1 Pourquoi l'entropie ?

En théorie de l'information, nous pouvons recenser un certain nombre de mesures de l'information contenue dans un signal [CT91, Gra90]. Une de ces mesures fondamentales est l'entropie.

Historiquement, l'entropie en théorie de l'information, aussi appelée entropie de Shannon car introduite par Claude Shannon en 1948 [Sha48], est une mesure plus large que l'entropie en thermodynamique et a été appelée ainsi par analogie avec cette notion de thermodynamique mesurant le désordre d'un système et possédant quelques propriétés identiques.

L'entropie d'une variable aléatoire est une mesure de la quantité moyenne d'incertitude de la variable aléatoire. Elle quantifie le nombre moyen de bits pour représenter une variable aléatoire. Pour un signal, l'entropie correspond à la quantité d'information contenue dans ce signal. En compression, elle correspond au nombre minimum de bits nécessaire pour coder un signal sans perte d'information. Plus un signal est redondant, plus son entropie est faible et l'entropie sera maximale pour une source dont toutes les valeurs sont équiprobables.

Dans le cas d'images, nous pouvons considérer que l'intensité lumineuse est une variable aléatoire et nous pouvons alors calculer l'entropie de la loi de probabilités de l'intensité lumineuse d'une image ou d'une portion d'image. Minimiser l'entropie d'une région conduit à une région la plus redondante possible en termes de valeurs d'intensité et donc à avoir une zone la plus homogène possible. L'entropie peut alors être considérée comme un critère d'homogénéité pour la région de l'image. Nous utiliserons le terme "homogène" dans le sens "proche d'une variable constante" et non pas sa définition probabiliste.

3.3.2 Estimation de l'entropie

Nous cherchons à estimer l'entropie de Shannon d'une densité de probabilité p :

$$H(p) = - \int_{\mathcal{S}} p(x) \log p(x) dx \quad (3.23)$$

où $\mathcal{S}$ est le support de p . L'entropie $H(p)$ est aussi parfois notée $H(X)$ avec X la variable aléatoire dont la densité de probabilité est p . Dans la définition de l'entropie le log est en général un log en base 2 afin de calculer des bits. Nous le noterons cependant log dans ce manuscrit.

Il existe de nombreux moyens d'estimer de façon non-paramétrique l'entropie d'un signal [BDGvdM97, WTP05]. Nous pouvons citer des méthodes d'estimation par substitution [AL76] qui consistent à remplacer la densité de probabilités p par une estimation $\hat{p}$ de celle-ci et des méthodes n'utilisant pas l'estimation des distributions des données, comme l'estimateur par espacement [Vas76] ou l'estimateur des k -ièmes plus proches voisins [KL87, GLMI05].

Nous choisissons d'utiliser l'estimateur de Ahmad-Lin qui repose sur un estimateur à noyaux de la densité p . C'est un estimateur qui donne de bonnes performances selon le choix du paramètre de lissage de l'estimateur à noyaux de la densité. L'estimateur de Ahmad-Lin de l'entropie de p s'écrit :

$$H(p) = \frac{1}{n} \int_{\mathcal{S}} -\log \hat{p}(x) dx . \quad (3.24)$$

avec n le nombre de points de l'échantillon.

Dans le contexte de la segmentation d'images, nous considérons donc la caractéristique $\mathbf{f}$ de l'image et p sa densité de probabilité sur une région Ω de l'image. L'expression de l'entropie de p sur la région Ω de l'image est donnée par :

$$H(\Omega) = \frac{1}{|\Omega|} \int_{\Omega} -\log \hat{p}(\mathbf{f}(\mathbf{x}), \Omega) d\mathbf{x} \quad (3.25)$$

où $\hat{p}$ est la densité estimée par la méthode à noyaux.

Nous obtenons ainsi une formulation continue de l'entropie, avec une densité de probabilités estimée par une méthode à noyaux (voir paragraphe 3.2.1). Par la suite et en

particulier dans les parties d'applications (parties 2 et 3), la densité de probabilité p sera toujours estimée par un estimateur à noyaux mais son estimation $\hat{p}$ sera cependant notée p pour une simplification d'écriture.

En général en traitement d'images, le critère utilisé est celui de la log-vraisemblance et la log-vraisemblance est une estimation de l'opposé de l'entropie sur toute l'image. Nous verrons que notre critère se rapproche de cette méthode lorsqu'il est utilisé en compétition de régions puisqu'intervient la probabilité d'appartenance à la région lors de l'utilisation de l'entropie conditionnelle. L'avantage de considérer l'entropie plutôt que le log-vraisemblance est qu'il est possible d'estimer l'entropie sans estimer la densité de probabilité. Ce critère d'entropie peut d'ailleurs être utilisé en compression d'images [AABG07].

3.3.3 Entropie jointe et entropie conditionnelle

Dans le cas où l'on voudrait prendre en compte plusieurs données, il faut faire intervenir la notion d'entropie jointe. L'entropie jointe entre les différentes variables correspond naturellement à la définition initiale de l'entropie en considérant les densités de probabilité conjointes. Si X et Y sont deux variables aléatoires dont la densité de probabilité conjointe s'écrit $p(x, y)$, l'entropie jointe de X et Y est donnée par :

$$H(X, Y) = - \int_S p(x, y) \log p(x, y) dx dy \quad (3.26)$$

avec S le support de p . Cette notion serait extensible à plus de deux variables en utilisant la densité de probabilités correspondante.

L'entropie jointe est reliée à l'entropie conditionnelle par la relation suivante :

$$H(X, Y) = H(X) + H(Y|X) \quad (3.27)$$

avec $H(Y|X)$ l'entropie conditionnelle de Y sachant X :

$$H(Y|X) = - \int_S p(y, x) \log p(y|x) dx dy \quad (3.28)$$

$$= - \int_{S_x} p(x) H(Y|X = x) dx \quad (3.29)$$

avec $p(y|x)$ la densité de probabilité de y sachant x et S_x le support de $p(x)$.

Nous utilisons le même estimateur que pour estimer l'entropie, à savoir l'estimateur de Ahmad-Lin.

D'autres mesures de l'information que l'entropie peuvent aussi être utilisées, il existe notamment des divergences qu'il peut être intéressant d'étudier.

3.4 DIVERGENCES

D'autres mesures de l'information peuvent être utilisées comme critère de segmentation. Tout comme l'entropie mesure l'incertitude d'une seule variable aléatoire, l'**information**

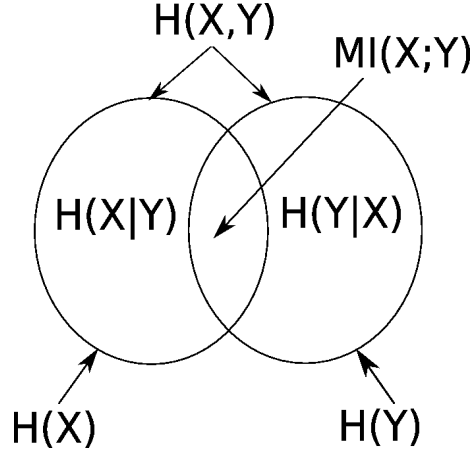


FIG. 3.5 – Relations entre entropie $H(\cdot)$, entropie conditionnelle $H(\cdot|\cdot)$ et information mutuelle $MI(\cdot;\cdot)$.

mutuelle (MI) mesure la réduction de cette incertitude due à une seconde variable aléatoire. C'est la mesure de la dépendance entre deux variables aléatoires. La figure 3.5 montre les relations entre ces différentes mesures.

La notion de distance entre mesures de l'information peut aussi être considérée. Par exemple, la distance entre une densité de probabilité et une densité de probabilité de référence.

3.4.1 Information mutuelle

L'information mutuelle de deux variables aléatoires mesure la dépendance statistique de ces variables. Si on note X et Y ces deux variables aléatoires, l'information mutuelle est définie par :

$$MI(X;Y) = \int p(x,y) \log \frac{p(x,y)}{p(x)p(y)} dx dy \quad (3.30)$$

Elle peut s'écrire en fonction de l'entropie :

$$\begin{aligned} MI(X;Y) &= H(X) - H(X/Y) = H(Y) - H(Y/X) \\ &= H(X) + H(Y) - H(X,Y) \end{aligned}$$

Nous utiliserons cette mesure dans le chapitre 6 en considérant l'information mutuelle entre l'intensité lumineuse et un label binaire.

3.4.2 Divergence de Kullback-Leibler

L'information mutuelle est en fait un cas particulier d'une mesure plus générale nommée divergence de Kullback-Leibler [KL51]. Cette divergence mesure une dissimilarité entre deux distributions p et q . Elle est définie par :

$$KL(p,q) = \int_S p(x) \log \frac{p(x)}{q(x)} dx \quad (3.31)$$

C'est une mesure qui est non-symétrique puisque $KL(p, q) \neq KL(q, p)$ et de plus, elle ne respecte pas l'inégalité triangulaire, mais elle est parfois appelée, par abus de langage, distance de Kullback-Leibler. En effet plus p et q sont proches, plus faible est cette mesure, et de plus $KL(p, q) = 0$ si et seulement si $p = q$, et $KL(p, q) > 0 \forall p, q$.

C'est une mesure qui joue un rôle important en théorie de l'information. Elle est souvent utilisée pour mesurer une dissimilarité entre la densité de probabilités des observations et la densité de probabilités du modèle choisi. Elle est aussi appelée *entropie relative*, *information relative*, *entropie croisée*, *fonction de discrimination* ...

En général, la distance de Kullback-Leibler est utilisée comme une mesure entre une distribution réelle p et un modèle de distribution q (par exemple une distribution gaussienne ou un autre modèle de la famille exponentielle) afin d'approximer la distribution réelle p . Le théorème de Sanov [San57] permet de donner une interprétation physique de cette mesure [CT91], à savoir une mesure de proximité entre une distribution empirique et une distribution. Une façon commune d'approximer la distribution p par un modèle de distribution consiste à trouver la distribution q , appartenant à la famille choisie, qui minimise la divergence de Kullback-Leibler entre p et q .

L'information mutuelle peut s'écrire en fonction de l'entropie relative :

$$MI(X; Y) = KL(p(x, y), p(x)p(y)) \quad (3.32)$$

La divergence de Kullback-Leibler peut s'écrire :

$$KL(p, q) = \int_S p(x) \log \frac{p(x)}{q(x)} dx \quad (3.33)$$

$$= \int_S p(x) \log p(x) dx - \int_S p(x) \log q(x) dx \quad (3.34)$$

$$= -H(p) - \int_S p(x) \log q(x) dx \quad (3.35)$$

avec $H(p)$ l'entropie de p .

Nous verrons dans le chapitre 7 comment utiliser cette distance pour la segmentation d'images en minimisant une distance entre une densité de référence et une densité courante.

3.5 CONCLUSION

Afin de nous permettre de prendre en compte les distributions des données les plus réalistes possible, nous nous intéressons à des critères issus de la théorie de l'information prenant en compte des distributions non-paramétriques. Pour cela nous devons donc considérer les différentes façons d'estimer les densités de probabilité des données. Nous nous attachons principalement à des méthodes non-paramétriques utilisant des noyaux pour estimer les densités de probabilité. Ainsi la méthode à noyau de Parzen nous paraît adaptée. Une des difficultés de ce type de méthode est le réglage de la largeur de bande, c'est-à-dire

de la variance du noyau utilisé. Des méthodes de substitution permettent de régler automatiquement en fonction des données. Une fois la distribution des données obtenue, nous présentons différentes mesures d'information. L'entropie permet de mesurer l'information contenue dans un signal et semble un critère pertinent pour une mesure d'homogénéité des données. D'autres mesures permettent de faire des mesures d'information entre plusieurs sources. L'entropie jointe et l'information mutuelle semblent donc adaptées lorsque l'on dispose de plusieurs canaux comme dans le cas d'images couleur. La divergence de Kullback-Leibler permet de mesurer une distance entre distributions et semble pertinente lorsque l'on dispose de données de référence.

Nous avons présenté la dérivation d'un critère général dépendant de la densité de probabilités de caractéristiques de l'image afin de parvenir à l'équation d'évolution du contour actif. Nous allons maintenant voir comment manipuler ces différentes mesures d'information dans le cadre de la segmentation d'images et de séquences vidéo par contours actifs.

Deuxième partie

Segmentation d'images

Introduction à la segmentation d'images

La segmentation d'images doit regrouper des pixels ayant des caractéristiques communes. Pour segmenter une image, la difficulté repose dans l'élaboration de critères définissant les objets à segmenter. Souvent des descripteurs statistiques de régions sont utilisés pour obtenir des régions ayant des propriétés statistiques particulières.

Il est courant d'utiliser une formulation statistique en traitement d'images. Nous pouvons notamment citer les travaux de Geman et Geman [GG84] et Besag [Bes86] pour des formulations discrètes, Mumford et Shah [MS89] et Zhu et Yuille [ZY96] pour des formulations continues. Pour la segmentation d'images par contours actifs, plusieurs auteurs ont utilisé une formulation probabiliste [TJW01, PD05, Cre06, MRGG06]. Nous pouvons noter l'utilisation du maximum de vraisemblance comme critère de segmentation [GR96, ZY96, SBFAZ00, PD02]. Pour la segmentation d'images en deux régions, de nombreux travaux étendent l'approche de Chan et Vese [CV01] tandis qu'une approche multiphase permet de segmenter plusieurs objets [ZCMO96, YTW99, VC02, BW04]. Les critères statistiques reposent principalement sur une hypothèse des distributions des données, notamment gaussienne [CV01, RD02]. Cette notion peut être étendue au cas vectoriel en utilisant comme tenseurs les matrices de covariance de variables ayant des distributions normales multivariées [WV04, LRDF04]. Des contours actifs d'ordre supérieur ont été utilisés pour la détection de réseaux routiers [RJZ06]. Une approche non-paramétrique, *i.e.* ne faisant aucune hypothèse sur les distributions des données peut aussi être utilisée [ABFJB03, KFY⁺05, COS06, LGF00]. Une récente revue de ces méthodes peut être trouvée dans [CRD07].

Dans cette partie, nous utilisons les critères statistiques présentés dans le chapitre 3 pour la segmentation d'images. Nous présentons d'abord l'utilisation de l'entropie comme critère de segmentation dans le cas d'images en niveaux de gris. Puis nous considérons le cas d'images en couleurs en utilisant l'entropie conjointe et l'information mutuelle. Nous présentons ensuite le cas d'une distance entre fonctions de probabilité comme critère de segmentation. Puis finalement nous combinons l'entropie avec un critère de distance à une forme de référence en utilisant l'exemple d'une référence obtenue par les *watersheds*.

Chapitre 5

Entropie

Dans ce chapitre, nous utilisons l'entropie comme critère de segmentation pour des régions dans des images en niveaux de gris. Nous considérons donc un critère basé sur l'entropie dans un schéma de contours actifs. Ce critère est minimisé suivant la méthode des gradients de forme et nous permet d'obtenir l'équation d'évolution du contour actif qui va faire tendre le critère vers son minimum. Nous appliquons ce formalisme à des images en niveaux de gris afin de segmenter des régions dont l'entropie est minimale.

5.1 ENTROPIE COMME CRITÈRE DE SEGMENTATION

Nous rappelons l'écriture de l'approximation de l'entropie que nous considérons :

$$H(\Omega) = -\frac{1}{|\Omega|} \int_{\Omega} \log p(\mathbf{f}(\mathbf{x}), \Omega) d\mathbf{x} \quad (5.1)$$

avec $\mathbf{f}$ la caractéristique de l'image considérée et $p(\mathbf{f}(\mathbf{x}), \Omega)$ sa densité de probabilité sachant que le pixel $\mathbf{x}$ est dans la région Ω .

Dans le cas d'images en niveaux de gris, nous considérons l'intensité lumineuse de l'image $I : \Omega_I \rightarrow \mathbb{R}$. Le critère d'entropie s'écrit donc :

$$H(\Omega) = -\frac{1}{|\Omega|} \int_{\Omega} \log p(I(\mathbf{x}), \Omega) d\mathbf{x} \quad (5.2)$$

Nous rappelons aussi que la densité de probabilité de cette intensité est estimée à l'aide d'un estimateur à noyaux de Parzen :

$$p(I(\mathbf{x}), \Omega) = \frac{1}{|\Omega|} \int_{\Omega} K_h(I(\mathbf{x}) - I(\hat{\mathbf{x}})) d\hat{\mathbf{x}} \quad (5.3)$$

où K_h est le noyau gaussien de cette estimation.

L'entropie peut alors être considérée comme une mesure d'homogénéité de l'intensité lumineuse sur une région. Elle a de plus l'intéressante propriété d'être invariante par translation, ce qui lui permet une certaine robustesse aux changements de luminosité moyenne. En cas de changement de luminosité multiplicatif, l'entropie n'est dans ce cas, pas invariante.

5.2 MINIMISATION DE L'ENTROPIE

Nous considérons que nous avons estimé les densités de probabilité des caractéristiques étudiées par la méthode des noyaux de Parzen. La minimisation de l'entropie demande le calcul de la dérivée eulérienne du critère. Nous nous servons des outils de dérivation présentés dans le chapitre 2 et nous utilisons pour cela la dérivation calculée dans la section 3.2.

Nous cherchons à dériver le critère :

$$H(\Omega) = -\frac{1}{|\Omega|} \int_{\Omega} \log p(I(\mathbf{x}), \Omega) d\mathbf{x} . \quad (5.4)$$

Notons $\mathcal{A}(\Omega) = \frac{1}{|\Omega|}$ et $\mathcal{B}(\Omega) = \int_{\Omega} -\log p(I(\mathbf{x}), \Omega) d\mathbf{x}$. Nous avons déjà calculé la dérivée eulérienne de $\mathcal{A}$ (voir (3.15)) :

$$d\mathcal{A}(\Omega, \mathbf{V}) = \frac{\int_{\Gamma} (\mathbf{V} \cdot \mathbf{N})}{|\Omega|^2} ds \quad (5.5)$$

et pour dériver $\mathcal{B}$, nous utilisons la formule (3.21) calculée dans la section 3.2 en utilisant $\varphi(r) = -\log(r)$. Cette formule s'écrivait :

$$dJ_r(\Omega, \mathbf{V}) = - \int_{\Gamma} (A(s, \Omega) + \varphi(p(I(s), \Omega))) (\mathbf{V} \cdot \mathbf{N}) ds \quad (5.6)$$

avec $A(s, \Omega)$ un terme venant de la dépendance du descripteur à la région :

$$A(s, \Omega) = -\frac{1}{|\Omega|} \int_{\Omega} \partial_1 \varphi(p(I(\mathbf{x}), \Omega)) \left[p(I(\mathbf{x}), \Omega) - K_h(I(\mathbf{x}) - I(s)) \right] d\mathbf{x}$$

Alors, nous avons $\mathcal{B}$ sous la forme :

$$d\mathcal{B}(\Omega, \mathbf{V}) = - \int_{\Gamma} \left(A(s, \Omega) + \varphi(p(I(s), \Omega)) \right) (\mathbf{V} \cdot \mathbf{N}) ds \quad (5.7)$$

avec

$$A(s, \Omega) = -\frac{1}{|\Omega|} \int_{\Omega} \partial_1 \varphi(p(I(\mathbf{x}), \Omega)) \left[p(I(\mathbf{x}), \Omega) - K_h(I(\mathbf{x}) - I(s)) \right] d\mathbf{x}$$

Ce qui donne, en remplaçant $\varphi(r)$ par $-\log r$:

$$\begin{aligned} d\mathcal{B}(\Omega, \mathbf{V}) &= - \int_{\Gamma} \left[-\frac{1}{|\Omega|} \int_{\Omega} \frac{-1}{p(I(\mathbf{x}), \Omega)} \left(p(I(\mathbf{x}), \Omega) - K_h(I(\mathbf{x}) - I(s)) \right) d\mathbf{x} \right. \\ &\quad \left. - \log p(I(s), \Omega) \right] (\mathbf{V} \cdot \mathbf{N}) ds \\ &= \int_{\Gamma} \left(-1 + \frac{1}{|\Omega|} \int_{\Omega} \frac{K_h(I(\mathbf{x}) - I(s))}{p(I(\mathbf{x}), \Omega)} d\mathbf{x} + \log p(I(s), \Omega) \right) (\mathbf{V} \cdot \mathbf{N}) ds \end{aligned} \quad (5.8)$$

En combinant (5.5) et (5.8), nous pouvons calculer la dérivée eulérienne de H :

$$\begin{aligned} dH(\Omega, \mathbf{V}) &= \frac{1}{|\Omega|^2} \int_{\Gamma} (\mathbf{V} \cdot \mathbf{N}) ds \int_{\Omega} -\log p(I(\mathbf{x}), \Omega) d\mathbf{x} \\ &\quad + \frac{1}{|\Omega|} \int_{\Gamma} \left(-1 + \frac{1}{|\Omega|} \int_{\Omega} \frac{K_h(I(\mathbf{x}) - I(s))}{p(I(\mathbf{x}), \Omega)} d\mathbf{x} + \log p(I(s), \Omega) \right) (\mathbf{V} \cdot \mathbf{N}) ds . \end{aligned} \quad (5.9)$$

Ce qui donne :

$$\begin{aligned} dH(\Omega, \mathbf{V}) &= \frac{1}{|\Omega|} H(\Omega) \int_{\Gamma} (\mathbf{V} \cdot \mathbf{N})_s \\ &+ \frac{1}{|\Omega|} \int_{\Gamma} \left(-1 + \frac{1}{|\Omega|} \int_{\Omega} \frac{K_h(I(\mathbf{x}) - I(s))}{p(I(\mathbf{x}), \Omega)} d\mathbf{x} + \log p(I(s), \Omega) \right) (\mathbf{V} \cdot \mathbf{N}) ds . \end{aligned} \quad (5.10)$$

En regroupant les intégrales, nous obtenons :

$$\begin{aligned} dH(\Omega, \mathbf{V}) &= \int_{\Gamma} \frac{1}{|\Omega|} \left(H(\Omega) + \frac{1}{|\Omega|} \int_{\Omega} \frac{K_h(I(\mathbf{x}) - I(s))}{p(I(\mathbf{x}), \Omega)} d\mathbf{x} \right. \\ &\quad \left. - 1 + \log p(I(s), \Omega) \right) (\mathbf{V} \cdot \mathbf{N}) ds . \end{aligned} \quad (5.11)$$

Ce qui peut s'écrire sous la forme :

$$dH(\Omega, \mathbf{V}) = -\frac{1}{|\Omega|} \int_{\Gamma} (A(s, \Omega) + \varphi(p(I(s), \Omega))) (\mathbf{V} \cdot \mathbf{N}) ds . \quad (5.12)$$

avec

$$A(s, \Omega) = -\left(H(\Omega) - 1 + \frac{1}{|\Omega|} \int_{\Omega} \frac{K_h(I(\mathbf{x}) - I(s))}{p(I(\mathbf{x}), \Omega)} d\mathbf{x} \right) .$$

De la dérivée (5.12), nous pouvons déduire l'équation d'évolution suivante pour le contour actif Γ :

$$\frac{\partial \Gamma}{\partial \tau}(\mathbf{s}) = \frac{1}{|\Omega|} (A(\mathbf{s}, \Omega) + \varphi(p(I(\mathbf{s}), \Omega))) \mathbf{N} \quad (5.13)$$

avec $\mathbf{s} = \Gamma(p, \tau)$, $\mathbf{N}$ la normale interne au contour actif et :

$$A(\mathbf{s}, \Omega) = -\left(H(\Omega) - 1 + \frac{1}{|\Omega|} \int_{\Omega} \frac{K_h(I(\mathbf{x}) - I(\mathbf{s}))}{p(I(\mathbf{x}), \Omega)} d\mathbf{x} \right)$$

et :

$$\varphi(p(I(\mathbf{s}), \Omega)) = -\log(p(I(\mathbf{s}), \Omega))$$

Nous utilisons cette équation d'évolution pour faire évoluer le contour actif vers l'objet d'intérêt que nous souhaitons segmenter.

5.3 RÉSULTATS EXPÉRIMENTAUX

Nous présentons dans cette partie les résultats expérimentaux obtenus en utilisant l'entropie comme critère d'homogénéité pour des images en niveau de gris.

Les expérimentations sont réalisées dans un cadre de contours actifs en compétition de régions [ZY96]. Notons Ω_{in} la région comprise à l'intérieur du contour Γ et notons Ω_{out} la région comprise à l'extérieur du contour. Le critère à minimiser dans ce cas est la somme

de l'entropie de l'intensité lumineuse à l'intérieur du contour et de l'entropie de l'intensité lumineuse à l'extérieur du contour :

$$J(\Gamma) = H(\Omega_{in}) + H(\Omega_{out}) \quad (5.14)$$

avec :

$$\begin{aligned} H(\Omega_{in}) &= -\frac{1}{|\Omega_{in}|} \int_{\Omega_{in}} \log p(I(\mathbf{x}), \Omega_{in}) d\mathbf{x} \\ H(\Omega_{out}) &= -\frac{1}{|\Omega_{out}|} \int_{\Omega_{out}} \log p(I(\mathbf{x}), \Omega_{out}) d\mathbf{x} . \end{aligned}$$

En minimisant le critère (5.14), nous obtenons l'équation d'évolution suivante pour le contour actif :

$$\begin{aligned} \frac{\partial \Gamma}{\partial \tau}(\mathbf{s}) &= \left[\frac{1}{|\Omega_{in}|} \left(A(\mathbf{s}, \Omega_{in}) - \log(p(I(\mathbf{s}), \Omega_{in})) \right) \right. \\ &\quad \left. - \frac{1}{|\Omega_{out}|} \left(A(\mathbf{s}, \Omega_{out}) - \log(p(I(\mathbf{s}), \Omega_{out})) \right) \right] \mathbf{N} \end{aligned} \quad (5.15)$$

avec :

$$A(\mathbf{s}, \Omega_i) = - \left(H(\Omega_i) - 1 + \frac{1}{|\Omega_i|} \int_{\Omega_i} \frac{K_h(I(\mathbf{x}) - I(\mathbf{s}))}{p(I(\mathbf{x}), \Omega_i)} d\mathbf{x} \right)$$

avec $i \in \{in, out\}$.

L'équation d'évolution en compétition de régions est obtenue en ajoutant les formules obtenues pour chacune des régions Ω_i en prenant en compte le fait que la normale intérieure $\mathbf{N}$ pour Ω_{out} est en sens contraire de celle de Ω_{in} , d'où le signe négatif entre les deux parties de l'équation. En ce faisant, nous cherchons à obtenir des régions homogènes, et pour l'objet et pour le fond.

Nous pouvons noter que cette équation comporte des termes locaux et des termes globaux. Les termes locaux sont ceux qui dépendent des coordonnées du pixel considéré, les termes globaux sont ceux qui sont égaux pour tous les pixels de la région. Ce sont les termes locaux qui permettent de faire évoluer chaque pixel différemment. Les termes prépondérants sont les deux termes $-\log(p(I(\mathbf{s}), \Omega_{in}))$ et $-\log(p(I(\mathbf{s}), \Omega_{out}))$ qui déterminent si un pixel a plus de probabilités d'appartenir à la région Ω_{in} ou à la région Ω_{out} . Si un pixel a une intensité proche de celle des pixels contenus à l'intérieur de la courbe, alors la probabilité $-\log(p(I(\mathbf{s}), \Omega_{in}))$ sera beaucoup plus élevée que $-\log(p(I(\mathbf{s}), \Omega_{out}))$ (même en étant normalisé par l'aire de la région) et l'algorithme fera évoluer la courbe de façon à inclure le pixel pour qu'il soit considéré comme faisant partie de l'objet d'intérêt.

Pour l'implémentation des contours actifs, nous utilisons une représentation explicite par B-splines. De ce fait il n'est pas nécessaire de rajouter de contrainte de régularisation sur le contour comme cela est fait fréquemment avec l'implémentation par lignes de niveaux, notamment pas de contrainte sur la longueur du contour. La régularisation est faite de façon intrinsèque par l'utilisation des B-splines en fonction du nombre de points d'échantillonnage utilisés. Les changements de topologie doivent par contre, être gérés de façon spécifique par l'implémentation.

5.3.1 Images de textures

Nous avons créé des images composites, formées de textures de Brodatz [Bro66] en niveaux de gris, de taille 256×256 pixels. Ces images sont donc composées d'un objet quelconque (pas forcément convexe) dans une première texture, sur un fond composé d'une seconde texture. Les images considérées sont des textures de métaux ou bien de tissus et ont une intensité quelconque, qui ne suit aucune loi spécifique. Nous observons les résultats de segmentation en considérant le critère d'entropie sur l'intensité lumineuse et nous comparons ces résultats à ceux obtenus avec un critère qui fait une supposition gaussienne sur les distributions de l'intensité.

La première image composite est présentée sur la figure 5.1. Les histogrammes des deux textures sont présentés également. Nous pouvons remarquer que les distributions de l'intensité lumineuse de ces deux textures sont assez proches d'une gaussienne, ou plus généralement d'une distribution paramétrique. L'initialisation est un cercle de rayon 110 pixels, centré dans l'image. La spline représentant le contour comporte 128 points d'échantillonnage. Le pas d'évolution est de 0.01 et le noyau de Parzen est une gaussienne de moyenne nulle et de variance 3. Les résultats de segmentation obtenus avec un critère d'entropie utilisant des distributions non-paramétriques et avec un critère utilisant des distributions paramétriques sont présentés sur la figure 5.2. Le critère utilisant des distributions paramétriques est une fonction de la variance de l'intensité lumineuse.

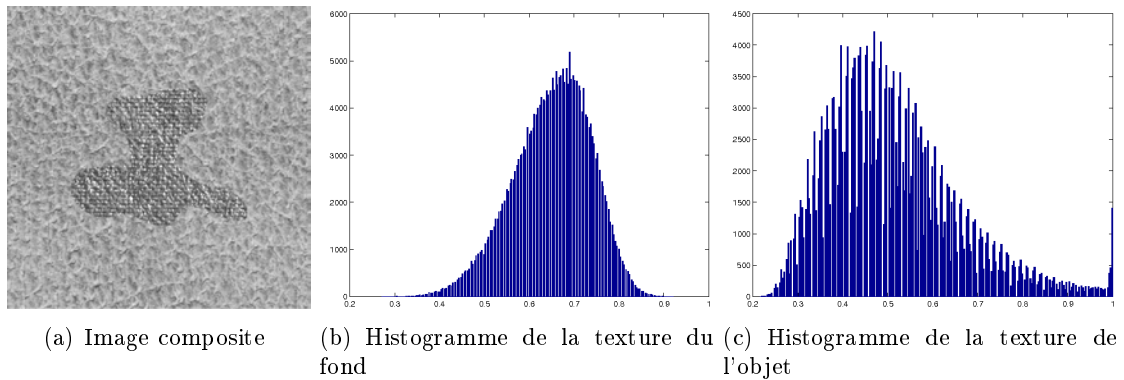


FIG. 5.1 – Image composée de textures et histogrammes des textures

Considérons une distribution gaussienne p de l'intensité de moyenne μ et de variance σ^2 . Celle-ci s'écrit alors :

$$p(\mathbf{x}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\mathbf{x} - \mu)^2}{2\sigma^2}\right)$$

Il est alors facile de montrer que l'entropie d'une variable suivant cette distribution dépend

de la variance [GD05] :

$$\begin{aligned}
H(X) &= - \int p(\mathbf{x}) \log p(\mathbf{x}) d\mathbf{x} \\
&= - \int \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\mathbf{x}-\mu)^2}{2\sigma^2}\right) \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\mathbf{x}-\mu)^2}{2\sigma^2}\right)\right) d\mathbf{x} \\
&= -\log \frac{1}{\sqrt{2\pi\sigma^2}} \int p(\mathbf{x}) d\mathbf{x} + \frac{1}{2\sigma^2} \int \frac{1}{\sqrt{2\pi\sigma^2}} (\mathbf{x}-\mu)^2 \exp\left(-\frac{(\mathbf{x}-\mu)^2}{2\sigma^2}\right) d\mathbf{x} \\
&= \frac{1}{2} \log(2\pi\sigma^2) + \frac{1}{2\sigma^2} E_p(x-\mu)^2
\end{aligned} \tag{5.16}$$

or on reconnaît l'écriture de l'espérance $\int (\mathbf{x}-\mu)^2 p(\mathbf{x})$. L'espérance de $(x-\mu)^2$ est donnée par :

$$E_p(x-\mu)^2 = \sigma^2 \tag{5.17}$$

Donc l'entropie s'écrit :

$$\begin{aligned}
H(X) &= \frac{1}{2} \log(2\pi\sigma^2) + \frac{1}{2} \\
&= \frac{1}{2} \log(2\pi\sigma^2 e) .
\end{aligned} \tag{5.18}$$

Le critère d'entropie est donc comparable à un critère dépendant de la variance dans le cas de distributions gaussiennes.

Nous pouvons remarquer sur la figure 5.2 que les résultats obtenus avec les deux méthodes donnent une segmentation convenable. Cela s'explique par le fait que les histogrammes des deux textures présentées sur la figure 5.1 se rapprochent d'une gaussienne et dans ce cas l'hypothèse de distribution gaussienne semble presque respectée et la segmentation est correcte en utilisant cette hypothèse de distribution paramétrique. Donc, dans un cas où les distributions se rapprochent de gaussiennes, le critère d'entropie marche aussi bien qu'un critère dépendant de la variance de la distribution de l'intensité.

Observons à présent une seconde image composite sur la figure 5.3. Cette image est toujours composée de deux textures dont les histogrammes sont affichés également sur la même figure. Cette fois-ci, l'histogramme de l'objet se rapproche toujours d'une gaussienne (il s'agit de la même texture que sur la figure 5.1) mais l'histogramme de la texture du fond est quelconque et très éloignée d'une gaussienne. Dans ce cas, l'hypothèse de distribution gaussienne n'est pas du tout respectée. L'initialisation est un cercle de rayon 70 pixels, situé en (158, 128). La spline initiale comporte 128 points d'échantillonnage. Le pas d'évolution est de 0.01 et le noyau de Parzen est une gaussienne de moyenne nulle et de variance 3. Les résultats de segmentation sont présentés sur la figure 5.4. Nous pouvons remarquer qu'avec le critère d'entropie la segmentation est correcte tandis qu'avec le critère dépendant de la variance de la distribution, la segmentation échoue et le contour reste bloqué sur la texture du fond. En effet, dans ce cas l'hypothèse de distribution gaussienne sous-entendu quand on utilise un critère dépendant de la variance est trop forte et n'est pas respectée, tandis qu'un critère comme l'entropie qui ne fait aucune hypothèse sur les distributions des données ne rencontre pas de telles difficultés et arrive à segmenter l'objet d'intérêt.

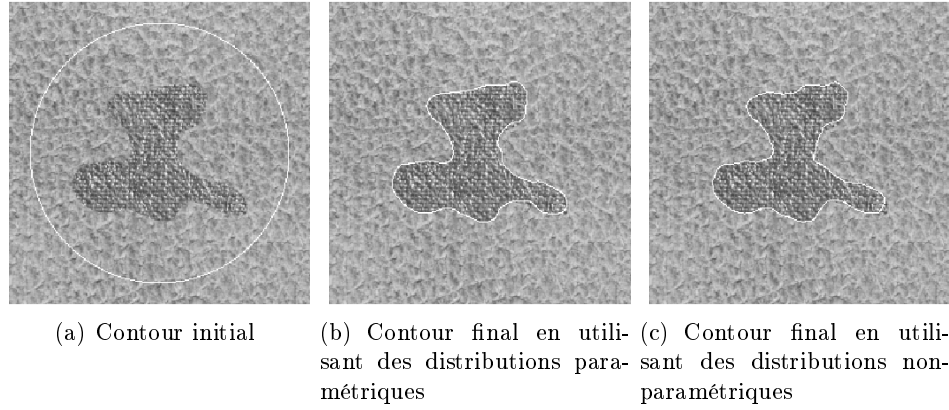


FIG. 5.2 – Courbe initiale et segmentation finale en utilisant un critère basé sur des distributions paramétriques et un critère basé sur des distributions non-paramétriques

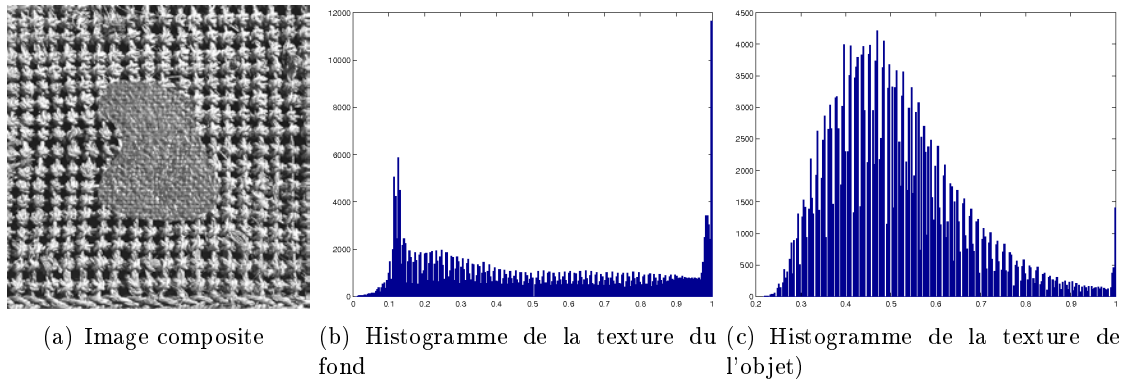


FIG. 5.3 – Image composée de textures et histogrammes des textures

5.3.2 Images médicales en niveau de gris

Après avoir testé notre critère d'entropie sur des images synthétiques de textures, testons-le sur des images médicales en niveau de gris. Il s'agit ici de segmenter des images servant au diagnostic d'ostéoporose. Ces images sont générées à partir de plusieurs images, de modalités différentes et sont de ce fait très bruitées. Elles ont une taille de 512×512 pixels. Nous pouvons remarquer que nous observons sur ces images des zones osseuses et des zones musculaires. Le but est de segmenter la zone d'intérêt définie par le médecin pour lui permettre, par calculs, d'affiner son diagnostic. La segmentation doit donc être précise autour de la zone d'intérêt et ne doit pas être perturbée par le bruit de l'image.

La figure 5.5 présente l'évolution de la segmentation au fur et à mesure des itérations ainsi que les histogrammes de l'objet (région comprise à l'intérieur du contour) et du fond (région comprise à l'extérieur du contour). L'initialisation choisie pour le contour actif est un cercle au centre de l'image de rayon 210 pixels. La spline initiale comporte 256 points

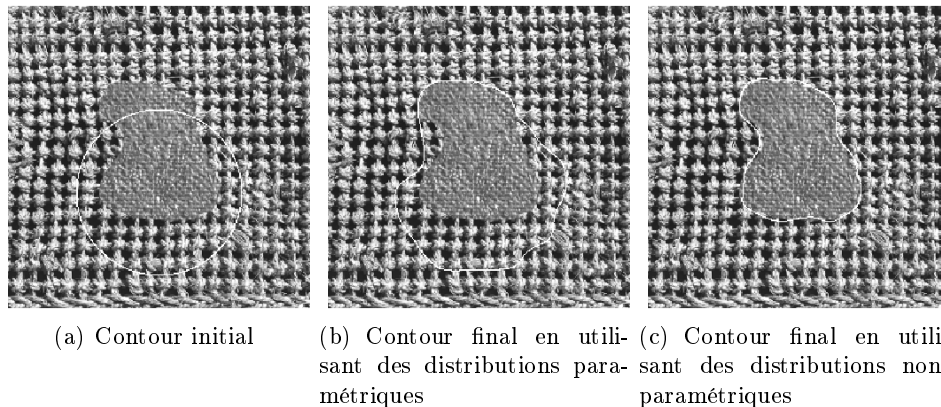


FIG. 5.4 – Courbe initiale et segmentation finale en utilisant un critère basé sur des distributions paramétriques et un critère basé sur des distributions non-paramétriques

d'échantillonnage et va évoluer vers la zone d'intérêt. Nous pouvons observer que le contour actif donne une segmentation lisse, qui ne s'accroche pas au bruit grâce à l'implémentation par B-splines et que la segmentation finale est précise, même aux endroits de forte courbure. L'évolution des histogrammes permet de constater que sur ce genre d'images, l'hypothèse gaussienne communément choisie pourrait être valide car les histogrammes finaux de l'objet et du fond se rapprochent de gaussiennes. Nous avons d'ailleurs comparé les histogrammes obtenus au fur et à mesure des itérations en minimisant l'entropie et en minimisant un critère dépendant de la variance, où les valeurs de variance et de moyenne des gaussiennes sont estimés à chaque itération. Cela est présenté sur la figure 5.6. Nous observons qu'à convergence de l'algorithme, les gaussiennes estimées sont assez proches des histogrammes réellement obtenus en minimisant l'entropie.

5.4 CONCLUSION

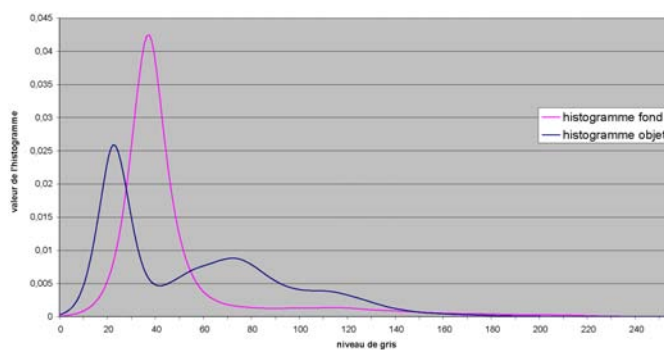
Nous avons présenté un critère de segmentation d'images qui repose sur l'entropie de l'intensité lumineuse de l'image. L'entropie mesurant la quantité d'information contenue dans un signal, minimiser l'entropie de l'intensité d'une région revient à chercher la région la plus homogène possible en terme d'intensité. Cependant, l'utilisation de ce critère permet une certaine variabilité de l'intensité lumineuse, d'autant plus quand il est utilisé en compétition de régions entre l'objet et le fond. Au minimum de la fonctionnelle, les régions sont telles que rajouter une portion du fond à l'objet (ou l'inverse) ferait augmenter l'entropie totale. Ce qui signifie par exemple que l'on ajouterait à la *pdf* de l'objet de nouvelles valeurs d'intensité, qui rajouteraient un mode à la *pdf* et feraient augmenter son entropie.

Nous avons vu dans les expériences que dans le cas de données dont la distribution ne suit pas une loi gaussienne, l'utilisation de distributions non-paramétrique et d'un critère comme l'entropie permet d'obtenir des segmentations correctes. En effet, le fait que les distributions des données soient quelconques ne perturbe pas l'algorithme, au contraire de

critères basés sur une hypothèse gaussienne pour les distributions, comme la variance. Mais une homogénéité en terme d'intensité n'apparaît pas suffisante pour des zones très texturées, il faudrait rajouter de l'information sur la structure des pixels afin de tenir compte que la localisation des pixels dans l'espace et non pas uniquement de leur intensité. Cela permettrait ainsi de segmenter correctement des zones homogènes et des zones texturées.



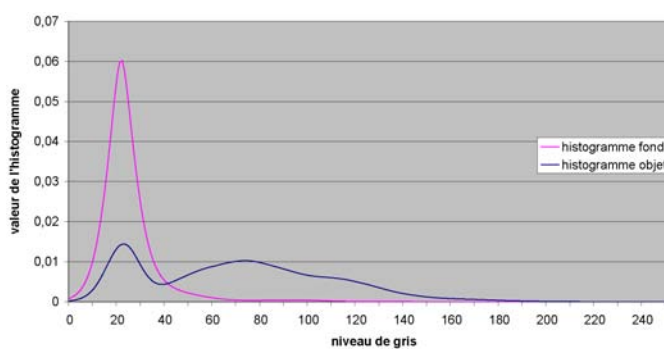
(a) Courbe initiale



(b) Histogrammes initiaux



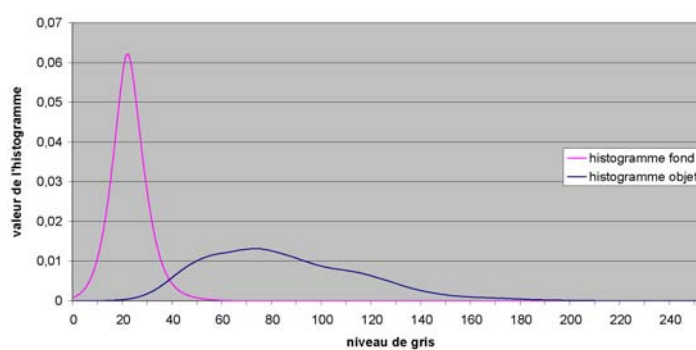
(c) Itération 100



(d) Itération 100



(e) Segmentation finale

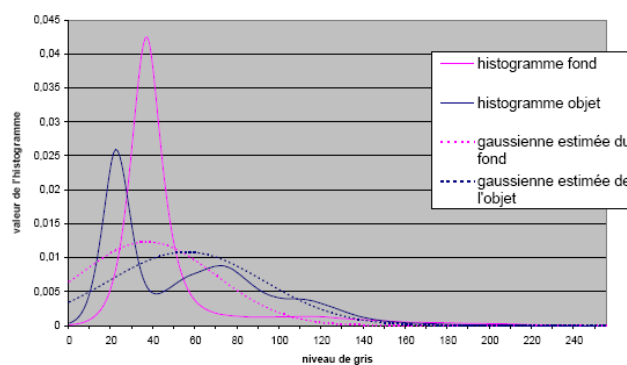


(f) Histogrammes finaux

FIG. 5.5 – Evolution de la segmentation et des histogrammes en minimisant l'entropie (histogramme de l'objet en bleu, histogramme du fond en rose)



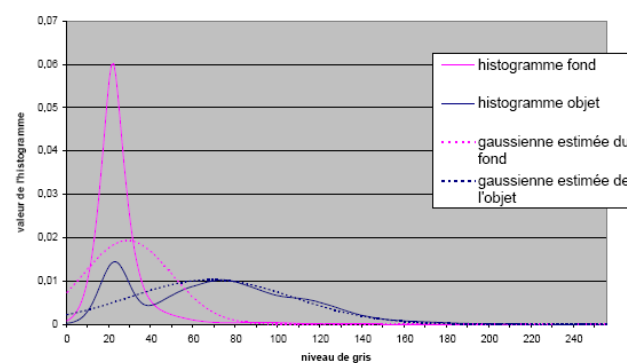
(a) Courbe initiale



(b) Histogrammes initiaux et gaussiennes estimées



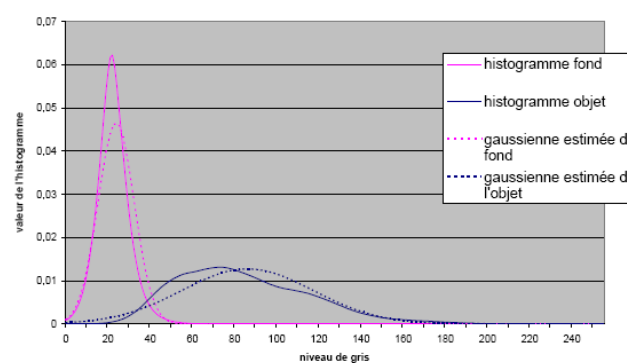
(c) Itération 100



(d) Itération 100



(e) Segmentation finale



(f) Histogrammes finaux et gaussiennes estimées

FIG. 5.6 – Comparaison des histogrammes obtenus en minimisant l'entropie et des gaussiennes estimées en minimisant une fonction de la variance. (histogramme de l'objet en bleu, histogramme du fond en rose, histogrammes en ligne pleine, gaussiennes en pointillés)

Entropie jointe et information mutuelle

Dans ce chapitre nous étendons le principe de segmentation vue au chapitre 5 pour le cas de données vectorielles. Nous appliquons la notion d'entropie aux images comportant plusieurs canaux comme les images en couleurs. Pour cela, nous considérons l'entropie jointe des canaux couleurs comme critère de segmentation. Nous cherchons ensuite à minimiser ce critère afin d'obtenir l'équation d'évolution du contour actif. Et pour tester l'efficacité de ce critère, nous l'appliquons à de la segmentation de visages. Nous présentons ensuite un critère basé sur l'information mutuelle afin de segmenter des zones plus ou moins homogènes dans des images couleur.

6.1 ENTROPIE JOINTE COMME CRITÈRE DE SEGMENTATION

Nous considérons des images comportant plusieurs canaux, comme par exemple les images en couleurs. Cette notion pourrait être également appliquée à différentes modalités d'images, comme par exemple des modalités de PET-SCAN et d'IRM dans le cas d'images médicales, ou des données multispectrales dans le cas d'images astrophysiques.

Notons $\mathbf{I} : \Omega_I \rightarrow \mathbb{R}^m$, l'intensité lumineuse de l'image avec m la dimension de l'espace des canaux de l'image, pour une image en couleurs $m = 3$ généralement. Nous considérons l'entropie sur une région de l'image Ω :

$$H(\Omega) = -\frac{1}{|\Omega|} \int_{\Omega} \log p(\mathbf{I}(\mathbf{x}), \Omega) d\mathbf{x} \quad (6.1)$$

avec $p(\mathbf{I}(\mathbf{x}))$ la densité de probabilité conjointe des différents canaux de $\mathbf{I}$ au pixel $\mathbf{x}$.

Nous utilisons ce critère de la même façon que dans le chapitre 5, juste en remplaçant la densité de probabilité par une densité de probabilité conjointe. Les densités de probabilité sont difficiles à estimer avec une méthode à noyaux quand on augmente la dimension car ces méthodes sont peu adaptées aux grandes dimensions. D'autres méthodes, basées notamment sur les k -plus proches voisins permettent d'estimer des densités de probabilité en grande dimension de façon plus adaptée [Fuk90].

La densité de probabilité conjointe est ainsi calculée de la façon suivante :

$$p(\mathbf{I}(\mathbf{x}), \Omega) = \frac{1}{|\Omega|} \int_{\Omega} K_h(\mathbf{I}(\mathbf{x}) - \mathbf{I}(\hat{\mathbf{x}})) d\mathbf{x} \quad (6.2)$$

avec K_h un noyau gaussien de dimension m , de moyenne nulle et de largeur de bande h :

$$K_h(u) = \frac{1}{(2\pi h)^{m/2}} \exp\left(-\frac{|u|^2}{2h^2}\right) \quad (6.3)$$

6.1.1 Minimisation de l'entropie jointe

La minimisation du critère (6.1) permet de parvenir à l'équation d'évolution du contour actif. En reprenant les équations obtenues au chapitre précédent nous obtenons la dérivée eulérienne suivante pour le critère (6.1) :

$$dH(\Omega, \mathbf{V}) = - \int_{\Gamma} \frac{1}{|\Omega|} (A(s, \Omega) + \varphi(p(\mathbf{I}(s), \Omega))) (\mathbf{V} \cdot \mathbf{N}) ds \quad (6.4)$$

avec

$$A(s, \Omega) = - \left(H(\Omega) - 1 + \frac{1}{|\Omega|} \int_{\Omega} \frac{K_h(\mathbf{I}(\mathbf{x}) - \mathbf{I}(s))}{p(\mathbf{I}(\mathbf{x}), \Omega)} d\mathbf{x} \right).$$

De cette dérivée, nous pouvons en déduire l'équation d'évolution du contour actif Γ :

$$\frac{\partial \Gamma}{\partial \tau}(\mathbf{s}) = \frac{1}{|\Omega|} (A(\mathbf{s}, \Omega) - \log(p(\mathbf{I}(\mathbf{s}), \Omega))) \mathbf{N} \quad (6.5)$$

avec $\mathbf{s} = \Gamma(p, \tau)$, $\mathbf{N}$ la normale interne au contour actif et :

$$A(\mathbf{s}, \Omega) = - \left(H(\Omega) - 1 + \frac{1}{|\Omega|} \int_{\Omega} \frac{K_h(\mathbf{I}(\mathbf{x}) - \mathbf{I}(\mathbf{s}))}{p(\mathbf{I}(\mathbf{x}), \Omega)} d\mathbf{x} \right).$$

L'équation d'évolution du contour actif étant obtenue, nous allons introduire cette équation dans un algorithme de contours actifs afin de segmenter des zones plus ou moins homogènes dans des images en couleur.

6.1.2 Résultats expérimentaux

Nous implémentons l'équation d'évolution (6.5) afin de segmenter des régions dans des images en couleurs. Pour cela, nous considérons des images dans l'espace de couleurs YUV où Y représente la luminance et U et V deux chrominances. Cet espace de couleurs est obtenu après conversion linéaire de l'espace couleurs RGB. Dans les expérimentations, nous ne considérons que deux des trois canaux, à savoir la luminance Y et la chrominance U pour des raisons de coût calculs. De plus, considérer les trois composantes couleur n'améliore pas de façon significative les résultats. Ces expérimentations sont effectuées dans un schéma de compétition de régions entre l'objet et le fond. Il s'agit donc de trouver la partition de l'image qui minimise à la fois l'entropie de l'objet et celle du fond.

Nous testons tout d'abord notre algorithme sur une image de la séquence vidéo **Erik**. Dans cette image, nous cherchons à segmenter le visage du personnage. La figure 6.1 montre l'évolution du contour actif et de l'histogramme de l'objet (partie comprise à l'intérieur du

contour) au fur et à mesure des itérations. L'initialisation est un cercle de rayon 80 pixels, centré dans l'image. La spline représentant le cercle est composé de 64 points d'échantillonnage. Le pas d'évolution est de 1,5 et la variance du noyau gaussien de 3. Nous pouvons remarquer que quand le contour se rapproche du visage du personnage, l'histogramme se fait plus ramassé et dense. Il s'agit en effet de segmenter une région assez homogène et de ce fait en minimisant l'entropie de la région nous cherchons à obtenir des histogrammes les plus ramassés possibles.

Nous testons ensuite notre critère sur des images en couleurs, pour chercher à segmenter les objets contenus dans l'image. Cela est présenté sur les figures 6.2, 6.3 et 6.4. L'image de la fleur a une taille de 640×427 pixels. L'initialisation est un cercle de rayon 70 pixels, positionné sur la fleur. Pour l'image du rhinocéros dont la taille fait 360×288 pixels, l'initialisation est un cercle de rayon 85 pixels. Dans le cas de l'écureuil dont les dimensions sont 288×209 pixels, le cercle initial a un rayon de 40 pixels. Pour les trois images, la spline initiale comporte 60 points d'échantillonnage, le pas d'évolution est de 1. Dans ces trois exemples, les objets d'intérêt présentent une couleur relativement homogène (sauf la fleur qui comporte deux couleurs principales mais qui sont très différentes de celles du fond). L'entropie est donc utilisée comme un critère d'homogénéité. Nous observons que les segmentations de la fleur, du rhinocéros et de l'écureuil sont acceptables même s'il manque un bout de l'écureuil parce que c'est une zone à l'ombre et donc beaucoup plus sombre que le reste de la robe de l'animal.

6.2 INFORMATION MUTUELLE

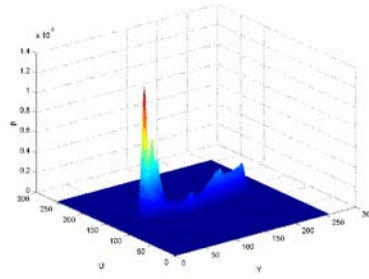
Nous présentons ensuite un critère de segmentation basé sur l'information mutuelle. L'information mutuelle a été beaucoup utilisée pour le recalage d'images, notamment dans l'imagerie médicale [WVA⁺96, MCV⁺97, VW97]. Nous cherchons à effectuer une segmentation de l'image en deux régions : l'objet et le fond. L'idée est ici de calculer l'information mutuelle entre l'intensité de l'image et un label qui détermine si l'on se situe à l'intérieur ou à l'extérieur de la courbe [KFY⁺05]. Notons L ce label binaire : $\Omega_I \rightarrow \{L_{in}, L_{out}\}$. Ω_{in} est la région de l'image comprise à l'intérieur de la courbe et Ω_{out} est la région comprise à l'extérieur de la courbe. Nous avons :

$$L(\mathbf{x}) = \begin{cases} L_{in} & \text{if } \mathbf{x} \in \Omega_{in} \\ L_{out} & \text{if } \mathbf{x} \in \Omega_{out} \end{cases} . \quad (6.6)$$

Nous cherchons à maximiser l'information mutuelle entre l'intensité de l'image $\mathbf{I}$ et le label L . En effet l'information mutuelle mesure la quantité d'information sur une variable quand on a la connaissance de l'autre variable. Le but est d'avoir au final la courbe qui sépare l'objet d'intérêt du fond. Les densités de probabilité de l'intensité de ces deux régions sont alors différentes. De ce fait plus le label peut déterminer si un pixel $\mathbf{x}$ appartient à Ω_{in} ou Ω_{out} , plus il contiendra d'information sur l'intensité de ce pixel. Maximiser l'information mutuelle MI entre $\mathbf{I}$ et L revient à chercher la partition $\{\Omega_{in}, \Omega_{out}\}$ de l'image qui collera au mieux à la segmentation correcte de l'image, c'est-à-dire celle qui sépare correctement les densités de probabilité.



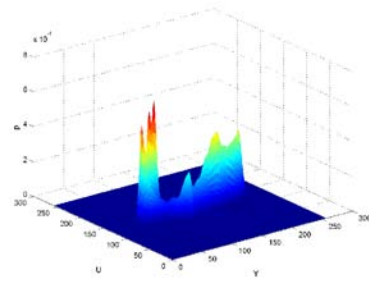
(a) Courbe initiale



(b) Histogramme initial de l'objet



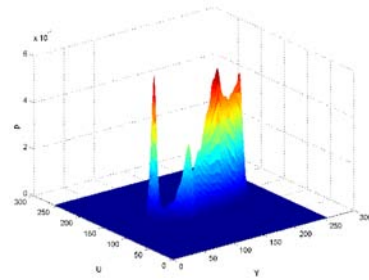
(c) Itération 100



(d) Itération 100



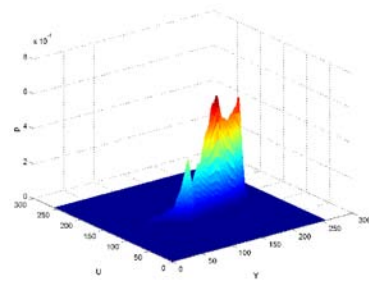
(e) Itération 300



(f) Itération 300



(g) Segmentation finale



(h) Histogramme final de l'objet

FIG. 6.1 – Evolution de la segmentation et des histogrammes en minimisant l'entropie jointe



FIG. 6.2 – Segmentation de la fleur en utilisant l'entropie jointe (canaux couleurs Y et U)

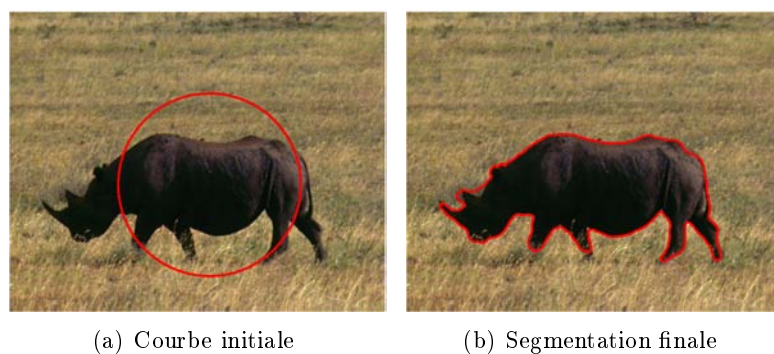


FIG. 6.3 – Segmentation du rhinocéros en utilisant l'entropie jointe (canaux couleurs Y et U)

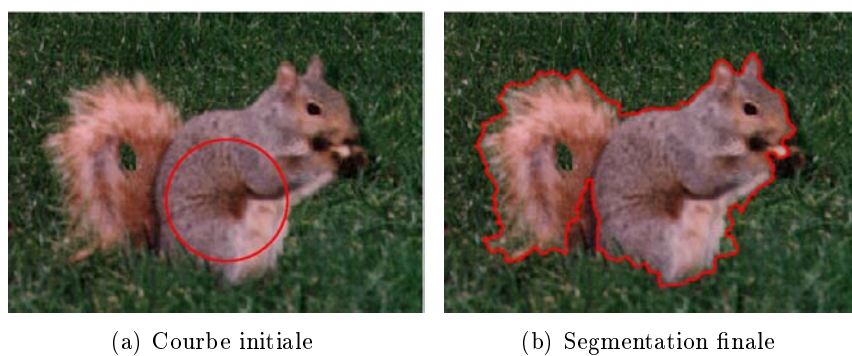


FIG. 6.4 – Segmentation de l'écureuil en utilisant l'entropie jointe (canaux couleurs Y et U)

Supposons que nous choisissons aléatoirement un point X dans Ω_I de façon à ce que X soit une localisation aléatoire uniformément distribuée dans le domaine de l'image. $L(X)$ est alors une variable aléatoire binaire qui dépend de la courbe Γ et $\mathbf{I}(X)$ est une variable aléatoire qui dépend des régions Ω_{in} et Ω_{out} . Nous considérons donc le critère suivant :

$$MI(\mathbf{I}(X); L(X)) = h(\mathbf{I}(X)) - h(\mathbf{I}(X)|L(X)) \quad (6.7)$$

où $h(Y)$ est l'entropie différentielle de la variable Y . Cette formulation peut être écrite comme :

$$\begin{aligned} MI(\mathbf{I}(X); L(X)) &= h(\mathbf{I}(X)) - h(\mathbf{I}(X)|L(X)) \\ &= h(\mathbf{I}(X)) - P(L(X) = L_{in})h(\mathbf{I}(X)|L(X) = L_{in}) \\ &\quad - P(L(X) = L_{out})h(\mathbf{I}(X)|L(X) = L_{out}) . \end{aligned} \quad (6.8)$$

Le critère à minimiser dans le cadre des contours actifs est alors :

$$\begin{aligned} J(\Gamma) &= -MI(\mathbf{I}(X); L(X)) \\ &= -h(\mathbf{I}(X)) + P(L(X) = L_{in})h(\mathbf{I}(X)|L(X) = L_{in}) \\ &\quad + P(L(X) = L_{out})h(\mathbf{I}(X)|L(X) = L_{out}) \end{aligned} \quad (6.9)$$

Or l'entropie de l'intensité de l'image sur le domaine est constante et ne dépend pas de la courbe donc nous pouvons supprimer ce terme. Les entropies conditionnelles $h(\mathbf{I}(X)|L(X) = L_{in})$ et $h(\mathbf{I}(X)|L(X) = L_{out})$ sont données par l'approximation de Ahmad-Lin :

$$h(\mathbf{I}(X)|L(X) = L_{in}) = \frac{1}{|\Omega_{in}|} \int_{\Omega_{in}} -\log p(\mathbf{I}(\mathbf{x}), \Omega_{in}) d\mathbf{x} \quad (6.10)$$

$$h(\mathbf{I}(X)|L(X) = L_{out}) = \frac{1}{|\Omega_{out}|} \int_{\Omega_{out}} -\log p(\mathbf{I}(\mathbf{x}), \Omega_{out}) d\mathbf{x} \quad (6.11)$$

Et les densités de probabilité sont estimées à l'aide de la méthode de Parzen. Le label $L(X)$ est une variable aléatoire binaire qui prend les valeurs L_{in} et L_{out} avec les probabilités $|\Omega_{in}|/|\Omega_I|$ et $|\Omega_{out}|/|\Omega_I|$, respectivement, puisque X est uniformément distribué dans le domaine de l'image. Le critère à minimiser devient alors :

$$\begin{aligned} J(\Gamma) &= \frac{|\Omega_{in}|}{|\Omega_I|} \frac{1}{|\Omega_{in}|} \int_{\Omega_{in}} -\log p(\mathbf{I}(\mathbf{x}), \Omega_{in}) d\mathbf{x} \\ &\quad + \frac{|\Omega_{out}|}{|\Omega_I|} \frac{1}{|\Omega_{out}|} \int_{\Omega_{out}} -\log p(\mathbf{I}(\mathbf{x}), \Omega_{out}) d\mathbf{x} \\ &= \frac{1}{|\Omega_I|} \int_{\Omega_{in}} -\log p(\mathbf{I}(\mathbf{x}), \Omega_{in}) d\mathbf{x} \\ &\quad + \frac{1}{|\Omega_I|} \int_{\Omega_{out}} -\log p(\mathbf{I}(\mathbf{x}), \Omega_{out}) d\mathbf{x} \end{aligned} \quad (6.12)$$

Comme le nombre de pixels sur le domaine de l'image Ω_I est constant, nous pouvons négliger cette constante et ce critère revient à minimiser la vraisemblance en compétition de régions.

6.2.1 Minimisation du critère

Nous utilisons les outils développés dans la section 2.3 pour dériver le critère (6.12) et parvenir à l'équation d'évolution du contour actif.

Nous obtenons la dérivée eulérienne dans la direction $\mathbf{V}$ suivante :

$$dJ(\Gamma, \mathbf{V}) = \int_{\Gamma} \left(\log(p(\mathbf{I}(s), \Omega_{in})) + \frac{1}{|\Omega_{in}|} \int_{\Omega_{in}} \frac{K_h(\mathbf{I}(\mathbf{x}) - \mathbf{I}(s))}{p(\mathbf{I}(\mathbf{x}), \Omega_{in})} d\mathbf{x} - 1 \right. \\ \left. - (\log(p(\mathbf{I}(s), \Omega_{out})) + \frac{1}{|\Omega_{out}|} \int_{\Omega_{out}} \frac{K_h(\mathbf{I}(\mathbf{x}) - \mathbf{I}(s))}{p(\mathbf{I}(\mathbf{x}), \Omega_{out})} d\mathbf{x} - 1) (\mathbf{V} \cdot \mathbf{N}) ds \right) \quad (6.13)$$

De cette dérivée nous pouvons déduire l'équation d'évolution du contour actif Γ :

$$\frac{\partial \Gamma}{\partial \tau}(\mathbf{s}) = \left(-\log(p(\mathbf{I}(s), \Omega_{in})) - \frac{1}{|\Omega_{in}|} \int_{\Omega_{in}} \frac{K_h(\mathbf{I}(\mathbf{x}) - \mathbf{I}(s))}{p(\mathbf{I}(\mathbf{x}), \Omega_{in})} d\mathbf{x} + 1 \right. \\ \left. - (-\log(p(\mathbf{I}(s), \Omega_{out})) - \frac{1}{|\Omega_{out}|} \int_{\Omega_{out}} \frac{K_h(\mathbf{I}(\mathbf{x}) - \mathbf{I}(s))}{p(\mathbf{I}(\mathbf{x}), \Omega_{out})} d\mathbf{x} + 1) \right) \mathbf{N} \quad (6.14)$$

Ce qui peut s'écrire :

$$\frac{\partial \Gamma}{\partial \tau}(\mathbf{s}) = \left(\log \frac{p(\mathbf{I}(\mathbf{s}), \Omega_{out})}{p(\mathbf{I}(\mathbf{s}), \Omega_{in})} - \frac{1}{|\Omega_{in}|} \int_{\Omega_{in}} \frac{K_h(\mathbf{I}(\mathbf{x}) - \mathbf{I}(\mathbf{s}))}{p(\mathbf{I}(\mathbf{x}), \Omega_{in})} d\mathbf{x} \right. \\ \left. + \frac{1}{|\Omega_{out}|} \int_{\Omega_{out}} \frac{K_h(\mathbf{I}(\mathbf{x}) - \mathbf{I}(\mathbf{s}))}{p(\mathbf{I}(\mathbf{x}), \Omega_{out})} d\mathbf{x} \right) \mathbf{N} \quad (6.15)$$

avec $\mathbf{s} = \Gamma(p, \tau)$.

Dans cette équation il n'y a que des termes locaux. Le terme prépondérant est le premier terme. En effet le terme $\log \frac{p(\mathbf{I}(s), \Omega_{out})}{p(\mathbf{I}(s), \Omega_{in})}$ permet, en comparant les probabilités de l'intensité, de déterminer si point $\mathbf{s}$ a plus de probabilité d'appartenir à la région Ω_{in} ou à la région Ω_{out} . Si $p(\mathbf{I}(\mathbf{s}), \Omega_{out})$ est supérieur à $p(\mathbf{I}(\mathbf{s}), \Omega_{in})$, cela va faire tendre le contour actif vers l'intérieur, sinon cela va le faire tendre vers l'extérieur. Cela signifie que la courbe englobera les pixels qui ont une plus forte probabilité d'appartenir à la région Ω_{in} qu'au fond Ω_{out} . Les deux autres termes locaux ont une faible influence sur l'équation d'évolution car leur valeur est très faible en comparaison du premier terme.

6.2.2 Résultats expérimentaux

Nous effectuons des tests sur des images de séquences vidéo, dans l'espace de couleurs YUV. Nous considérons les deux premiers canaux couleurs (Y et U) et nous appliquons notre méthode à la segmentation de visages. La figure 6.5 présente l'évolution du contour actif sur une image de la séquence **Foreman** tandis que la figure 6.6 montre l'évolution sur une image de la séquence **Erik**. Dans les deux séquences, l'initialisation est un cercle de rayon 80 pixels et la spline comporte 64 points d'échantillonnage. Le pas d'évolution est de 1.5 et la fenêtre de Parzen est une gaussienne de variance 3. La figure 6.7 montre le suivi de la segmentation le long de la séquence vidéo. Ce suivi est effectué en considérant comme initialisation sur une image le résultat de la segmentation de l'image précédente. Ainsi il est possible de suivre l'objet segmenté au fur et à mesure de la séquence vidéo.

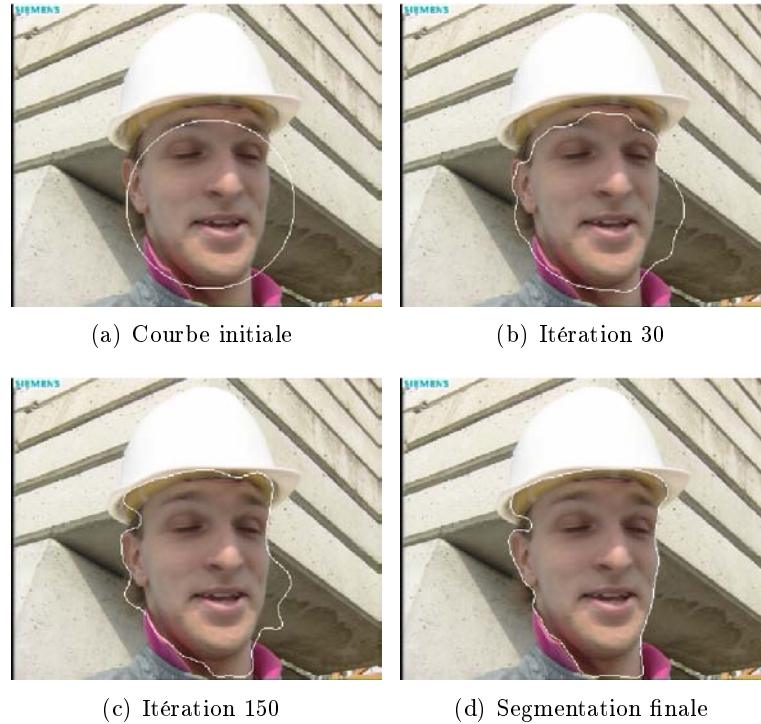


FIG. 6.5 – Evolution de la segmentation en minimisant l'information mutuelle

6.3 CONCLUSION

Pour la segmentation d'images couleur, nous avons considéré des critères basés sur l'entropie jointe et l'information mutuelle. L'entropie jointe entre les différents canaux couleur nous permet de segmenter des régions plus ou moins homogènes, comme cela est le cas dans le cas de la segmentation de visages. L'entropie permet une certaine variabilité dans les caractéristiques utilisées, d'autant plus quand elle est utilisée comme ici en compétition de régions. Le minimum de la fonctionnelle correspond alors à un compromis entre l'homogénéité du fond et l'homogénéité de l'objet, tout en assurant une certaine séparabilité des couleurs considérées. En effet, le minimum de la fonctionnelle correspond à la meilleure séparation possible des pdf des couleurs. Si on rajoute une portion du fond à l'objet, on rajoute des nouvelles couleurs à l'objet et on fait plus augmenter son entropie qu'on ne fait diminuer celle du fond (sinon nous ne serions pas au minimum de la fonctionnelle). Les régions ainsi obtenues ne sont donc pas forcément les plus homogènes possible, mais elles correspondent à un compromis en terme d'homogénéité et de séparation des couleurs entre l'objet et le fond.

En considérant l'information mutuelle entre l'intensité lumineuse et un label binaire d'appartenance à la région à l'intérieur de la courbe, nous nous ramenons à un critère de vraisemblance. Nous avons appliqué ces deux critères à la segmentation de visages dans des images couleur issues de séquences vidéo en considérant deux canaux couleur dans l'espace

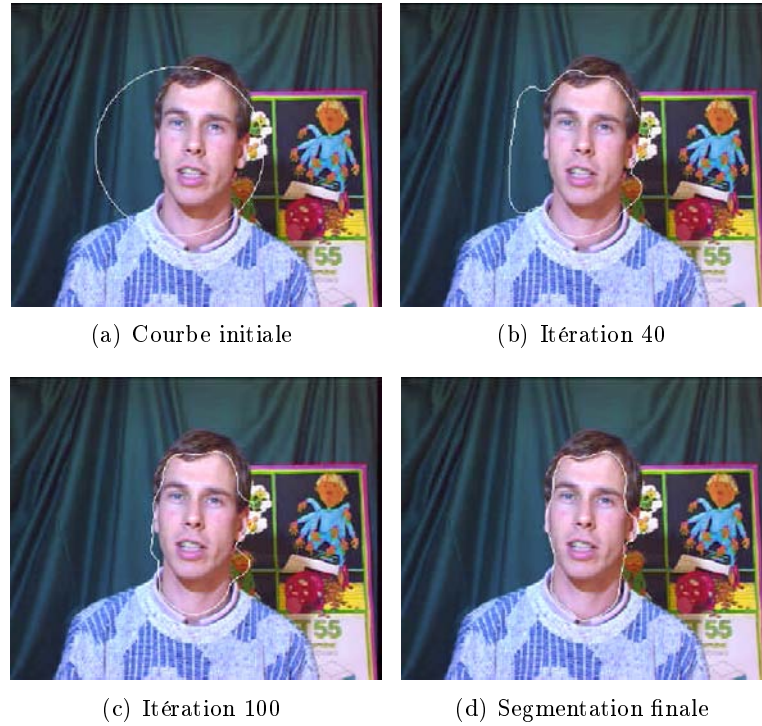


FIG. 6.6 – Evolution de la segmentation en minimisant l'information mutuelle

YUV. Le suivi de cette segmentation au cours de la séquence vidéo est aussi possible en prenant comme initialisation dans l'image suivante, le résultat de la segmentation de l'image précédente. La différence entre ces deux critères intervient au niveau de la normalisation par le nombre de pixels de la région. Dans le premier cas on cherche à minimiser l'entropie jointe de la région, dans le second cas, l'entropie jointe multipliée par le cardinal de la région. Il semblerait que cette différence intervienne au niveau de la taille de la région segmentée. Il serait intéressant de déterminer dans quel cas utiliser l'un ou l'autre des critères selon l'objet à segmenter.



FIG. 6.7 – Segmentation au cours de la séquence vidéo

Chapitre 7

Divergence de Kullback-Leibler

Dans ce chapitre nous proposons un critère de segmentation basé sur une distance entre distributions pour la segmentation d'images. Nous présentons tout d'abord le critère dépendant de la distance puis nous étudions ensuite le cas de la divergence de Kullback-Leibler.

7.1 CRITÈRE BASÉ SUR UNE DISTANCE ENTRE DISTRIBUTIONS

Dans cette section, nous proposons une segmentation qui prend en compte un critère basé sur une distance entre distributions. Cela peut être justifié par le fait que l'on dispose d'une segmentation de référence et que l'on cherche à segmenter un objet d'intérêt dans une autre image ayant les mêmes propriétés en terme de couleurs ou de niveaux de gris. Etant donné une portion Ω de l'image $\mathbf{I}$ et $\mathbf{x}$ un pixel de la région Ω , la probabilité d'avoir $\mathbf{I}(\mathbf{x}) = \alpha$ est notée $p(\alpha)$:

$$p(\alpha) = Pr\{\mathbf{I}(\mathbf{x}) = \alpha\}$$

7.1.1 Définition de la distance entre distributions

Supposons que nous connaissons la fonction de densité de probabilité de $\mathbf{I}$ sur une région de référence Ω_{ref} . Notons $p_{ref} : \mathbb{R} \rightarrow [0, 1]$ cette fonction. Le but est de faire évoluer une région Ω de façon à ce que sa fonction de densité de probabilité se rapproche de celle de la région de référence Ω_{ref} . Nous cherchons donc à minimiser une distance de la fonction de densité de probabilité $p(\alpha, \Omega)$. Nous utilisons une fonction $\varphi : \mathbb{R}^+ \times \mathbb{R}^m \rightarrow \mathbb{R}^+$ pour calculer la distance entre les distributions. La distance entre deux distributions est définie de la façon suivante :

$$D(\Omega) = \int_{\mathbb{R}^m} \varphi(p(\alpha, \Omega), \alpha) \, d\alpha . \quad (7.1)$$

La fonction φ peut être par exemple la fonction quadratique :

$$\varphi(p(\alpha, \Omega), \alpha) = (p(\alpha, \Omega) - p_{ref}(\alpha))^2$$

ou bien telle que (7.1) soit le carré de la distance de Hellinger :

$$\varphi(p(\alpha, \Omega), \alpha) = (\sqrt{p(\alpha, \Omega)} - \sqrt{p_{ref}(\alpha)})^2$$

De nombreuses distances pourraient être choisies. Nous choisissons une mesure qui prenne en compte l'information comprise dans les données et c'est pour cela que nous nous intéressons à une mesure issue de la théorie de l'information. Dans les expérimentations, nous considérerons une version symétrisée de la divergence de Kullback-Leibler, nommée J-divergence [Jef46]. Généralement c'est une version non-symétrisée de la divergence de Kullback-Leibler qui est utilisée car même si elle n'est pas une distance, elle a des propriétés qui permettent de l'utiliser comme une mesure de proximité [Réf05, RG06]. Alors :

$$\varphi(p(\alpha, \Omega), \alpha) = \frac{1}{2} (p_{ref}(\alpha) \log \frac{p_{ref}(\alpha)}{p(\alpha, \Omega)} + p(\alpha, \Omega) \log \frac{p(\alpha, \Omega)}{p_{ref}(\alpha)})$$

Nous avons donc la fonctionnelle suivante :

$$\begin{aligned} D(\Omega) &= \int_{\mathbb{R}^m} \frac{1}{2} (p_{ref}(\alpha) \log \frac{p_{ref}(\alpha)}{p(\alpha, \Omega)} + p(\alpha, \Omega) \log \frac{p(\alpha, \Omega)}{p_{ref}(\alpha)}) d\alpha \\ &= \frac{1}{2} (KL(p_{ref}, p) + KL(p, p_{ref})) \end{aligned} \quad (7.2)$$

en notant $KL(., .)$ la distance de Kullback-Leibler. Nous pouvons analyser les deux termes intervenant dans la version symétrisée (voir à ce sujet [Min05]). En minimisant le premier terme dans l'équation 7.2, il s'agit de la distance habituellement utilisée afin de trouver une distribution p qui approxime au mieux la distribution p_{ref} . En utilisant cette distance, la distribution p va essayer de couvrir toute l'étendue de la distribution p_{ref} , et si $p_{ref} \geq 0$, alors $p \geq 0$. La minimisation d'un tel terme a l'avantage d'être assez stable numériquement. La minimisation de l'autre terme implique que la distribution p va se concentrer sur la partie prédominante de la distribution p_{ref} . Par exemple, si on cherche à se rapprocher d'une distribution p_{ref} contenant deux modes par une distribution p gaussienne, p va se rapprocher du mode prédominant de p_{ref} en excluant l'autre. Cette approche a l'inconvénient de fournir une distribution avec une variance trop petite mais selon les problèmes cela peut être intéressant. En choisissant une combinaison de ces deux termes nous faisons comme un "compromis" entre ces deux approches et surtout en symétrisant le critère, nous obtenons les propriétés d'une distance.

Cette fonctionnelle va nous permettre de faire évoluer le contour actif vers une région dont les distributions seront proches de celles de la région de référence.

7.1.2 Vers l'équation d'évolution du contour actif

Pour parvenir à l'équation d'évolution du contour actif, nous calculons la dérivée eulérienne dans la direction $\mathbf{V}$ de la distance $D(\Omega)$ définie en (7.1) en utilisant les outils de dérivation des gradients de forme.

La dérivée eulérienne de $D(\Omega)$ est donnée par :

$$dD(\Omega, \mathbf{V}) = \int_{\mathbb{R}^m} d\varphi(p(\alpha, \Omega), \alpha, \mathbf{V}) d\alpha \quad (7.3)$$

Nous définissons :

$$\varphi(p(\alpha, \Omega), \alpha) \stackrel{\text{def}}{=} f(G_1(\alpha, \Omega).G_2(\Omega)) = \varphi\left(\frac{G_1(\alpha, \Omega)}{G_2(\Omega)}, \alpha\right)$$

avec

$$G_1(\alpha, \Omega) = \int_{\Omega} K_h(\mathbf{I}(\mathbf{x}) - \alpha) \, d\mathbf{x}$$

$$G_2(\Omega) = |\Omega| = \int_{\Omega} d\mathbf{x}$$

La dérivée eulérienne de $\varphi(p(\alpha, \Omega), \alpha)$ est donnée par :

$$d\varphi(p(\alpha, \Omega), \alpha, \mathbf{V}) = df(\Omega, \mathbf{V}) = \frac{\partial f}{\partial G_1} dG_1(\alpha, \Omega, \mathbf{V}) + \frac{\partial f}{\partial G_2} dG_2(\alpha, \Omega, \mathbf{V})$$

Nous posons $h(\alpha, \mathbf{x}) = K_h(\mathbf{I}(\mathbf{x}) - \alpha)$ et nous calculons les dérivées eulériennes de G_1 et G_2 :

$$dG_1(\alpha, \Omega, \mathbf{V}) = \int_{\Omega} h'(\alpha, \mathbf{x}, \mathbf{V}) \, d\mathbf{x} - \int_{\Gamma} h(\alpha, s)(\mathbf{V} \cdot \mathbf{N}) \, ds$$

h ne dépend pas de Ω donc la dérivée eulérienne de G_1 vaut :

$$dG_1(\alpha, \Omega, \mathbf{V}) = - \int_{\Gamma} h(\alpha, s)(\mathbf{V} \cdot \mathbf{N}) \, ds$$

Pour G_2 , nous trouvons de la même façon :

$$dG_2(\alpha, \Omega, \mathbf{V}) = - \int_{\Gamma} (\mathbf{V} \cdot \mathbf{N}) \, ds$$

La dérivée de f par rapport à dG_1 est :

$$\begin{aligned} \frac{\partial f}{\partial G_1}(G_1(\alpha, \Omega), G_2(\Omega)) &= \frac{1}{G_2(\Omega)} \frac{\partial \varphi}{\partial G_1}\left(\frac{G_1(\alpha, \Omega)}{G_2(\Omega)}, \alpha\right) \\ &= \frac{1}{|\Omega|} \partial_1 \varphi(p(\alpha, \Omega), \alpha) \end{aligned}$$

avec $\partial_1 \varphi$ la dérivée partielle de φ par-rapport à la première variable.

La dérivée de f par rapport à dG_2 est :

$$\begin{aligned} \frac{\partial f}{\partial G_2}(G_1(\alpha, \Omega), G_2(\Omega)) &= \frac{-G_1(\alpha, \Omega)}{G_2(\Omega)^2} \frac{\partial \varphi}{\partial dG_2}\left(\frac{G_1(\alpha, \Omega)}{G_2(\Omega)}, \alpha\right) \\ &= -\frac{p(\alpha, \Omega)}{|\Omega|} \partial_1 \varphi(p(\alpha, \Omega), \alpha) \end{aligned}$$

Alors la dérivée eulérienne de f vaut :

$$df(\Omega, \mathbf{V}) = -\frac{\partial_1 \varphi(p(\alpha, \Omega), \alpha)}{|\Omega|} \int_{\Gamma} (K_h(\mathbf{I}(s) - \alpha) - p(\alpha, \Omega))(\mathbf{V} \cdot \mathbf{N}) \, ds$$

En remplaçant cette expression dans (7.3), nous avons l'expression suivante pour la dérivée eulérienne de $D(\Omega)$:

$$\begin{aligned} dD(\Omega, \mathbf{V}) &= - \int_{\Gamma} \frac{1}{|\Omega|} \left(\int_{\mathbb{R}^m} (\partial_1 \varphi(p(\alpha, \Omega), \alpha) K_h(\mathbf{I}(s) - \alpha)) \, d\alpha \right. \\ &\quad \left. - \int_{\mathbb{R}^m} \partial_1 \varphi(p(\alpha, \Omega), \alpha) p(\alpha, \Omega) \, d\alpha \right) (\mathbf{V} \cdot \mathbf{N}) \, ds \end{aligned} \quad (7.4)$$

Dans le cas de la divergence de Kullback-Leibler, la fonction φ s'écrit :

$$\varphi(p(\alpha, \Omega), \alpha) = \frac{1}{2} (p_{ref}(\alpha) \log \frac{p_{ref}(\alpha)}{p(\alpha, \Omega)} + p(\alpha, \Omega) \log \frac{p(\alpha, \Omega)}{p_{ref}(\alpha)})$$

Ce qui donne pour la dérivée :

$$\partial_1 \varphi(p(\alpha, \Omega), \alpha) = \frac{1}{2} \left(\frac{-p_{ref}(\alpha)}{p(\alpha, \Omega)} + \log \frac{p(\alpha, \Omega)}{p_{ref}(\alpha)} + 1 \right)$$

Dans ce cas, la dérivée eulérienne de $D(\Omega)$ s'écrit :

$$\begin{aligned} dD(\Omega, \mathbf{V}) = & - \int_{\Gamma} \frac{1}{|\Omega|} \left(\int_{\mathbb{R}^m} \left(\frac{1}{2} \left(\frac{-p_{ref}(\alpha)}{p(\alpha, \Omega)} + \log \frac{p(\alpha, \Omega)}{p_{ref}(\alpha)} + 1 \right) K_h(\mathbf{I}(s) - \alpha) \right) d\alpha \right. \\ & \left. - \int_{\mathbb{R}^m} p(\alpha, \Omega) \left(\frac{1}{2} \left(\frac{-p_{ref}(\alpha)}{p(\alpha, \Omega)} + \log \frac{p(\alpha, \Omega)}{p_{ref}(\alpha)} + 1 \right) \right) p(\alpha, \Omega) d\alpha \right) (\mathbf{V} \cdot \mathbf{N}) ds \end{aligned} \quad (7.5)$$

Nous utilisons le critère $D(\Omega)$ en compétition de régions, c'est-à-dire que nous cherchons à minimiser la distance entre la distribution de l'objet et celle de l'objet de référence et la distance entre la distribution du fond et celle du fond de référence.

De ce fait, la fonctionnelle à minimiser est :

$$J(\Gamma) = D(\Omega_{in}) + D(\Omega_{out}) \quad (7.6)$$

De la dérivée eulérienne (7.5), nous pouvons en déduire l'équation d'évolution du contour actif :

$$\begin{aligned} \frac{\partial \Gamma}{\partial \tau}(\mathbf{s}) = & \frac{1}{2|\Omega_{in}|} \left(\int_{\mathbb{R}^m} \left(\left(\frac{-p_{ref_{in}}(\alpha)}{p(\alpha, \Omega_{in})} + \log \frac{p(\alpha, \Omega_{in})}{p_{ref_{in}}(\alpha)} + 1 \right) K_h(\mathbf{I}(\mathbf{s}) - \alpha) \right) d\alpha \right. \\ & \left. - \int_{\mathbb{R}^m} p(\alpha, \Omega_{in}) \left(\left(\frac{-p_{ref_{in}}(\alpha)}{p(\alpha, \Omega_{in})} + \log \frac{p(\alpha, \Omega_{in})}{p_{ref_{in}}(\alpha)} + 1 \right) \right) p(\alpha, \Omega_{in}) d\alpha \right) \mathbf{N} \\ & - \frac{1}{2|\Omega_{out}|} \left(\int_{\mathbb{R}^m} \left(\left(\frac{-p_{ref_{out}}(\alpha)}{p(\alpha, \Omega_{out})} + \log \frac{p(\alpha, \Omega_{out})}{p_{ref_{out}}(\alpha)} + 1 \right) K_h(\mathbf{I}(\mathbf{s}) - \alpha) \right) d\alpha \right. \\ & \left. - \int_{\mathbb{R}^m} p(\alpha, \Omega_{out}) \left(\left(\frac{-p_{ref_{out}}(\alpha)}{p(\alpha, \Omega_{out})} + \log \frac{p(\alpha, \Omega_{out})}{p_{ref_{out}}(\alpha)} + 1 \right) \right) p(\alpha, \Omega_{out}) d\alpha \right) \mathbf{N} \end{aligned} \quad (7.7)$$

avec $p_{ref_{in}}$ et $p_{ref_{out}}$ les densité de probabilités de référence de l'objet et du fond, respectivement.

Cette équation d'évolution permet de faire évoluer le contour actif vers une région dont les distributions seront proches de celles de référence. Etudions la première partie de l'équation qui concerne Ω_{in} . Dans cette équation le seul terme local est la première intégrale qui correspond à une convolution par le noyau K_h . Ce terme permet de comparer la valeur de la densité de référence à celle estimée pour la valeur $\mathbf{I}(\mathbf{s})$. Si cette convolution est positive, cela signifie que la densité de probabilité estimée de $\mathbf{I}(\mathbf{s})$ est plus forte que celle de référence et l'algorithme va exclure ce point de l'objet. Le même raisonnement peut être fait pour Ω_{out} .

7.2 RÉSULTATS EXPÉRIMENTAUX

Nous utilisons l'équation d'évolution (7.7) afin de segmenter une région d'intérêt en connaissant une segmentation de référence. Dans le cas d'une séquence vidéo, il peut s'agir de la connaissance de la segmentation d'une image de la séquence pour en déduire la segmentation de l'image suivante ou bien d'une autre image de la séquence vidéo. Nous cherchons donc à segmenter une région qui comporte les mêmes caractéristiques en terme de couleurs à la région de référence. Cela s'avère efficace notamment quand l'objet à segmenter subit des déformations de forme tout en préservant les caractéristiques statistiques de l'objet.

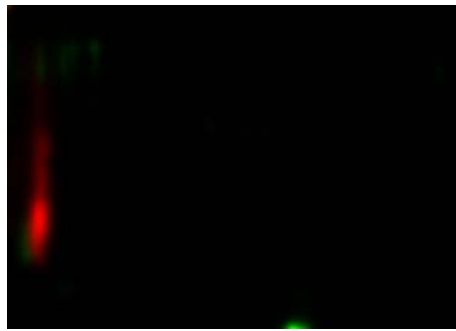
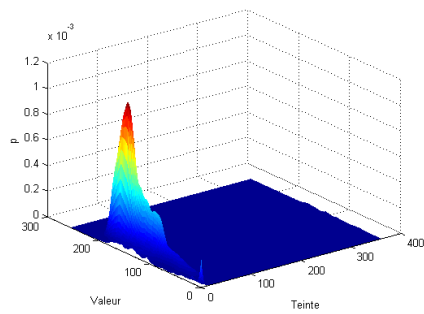
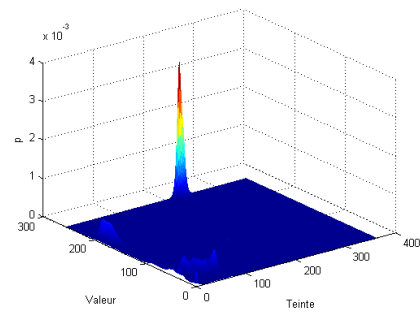
Nous présentons des résultats sur les séquences d'images **Akiyo** du groupe MPEG et **Phone** de France Telecom.

Nous cherchons à segmenter le visage de la présentatrice dans la séquence **Akiyo**. Nous faisons l'hypothèse que nous disposons d'une segmentation de référence, ici sur l'image 20 de la séquence. Les densités de probabilité de référence de l'objet d'intérêt (ici le visage) et du fond sont calculés et représentés sur la figure 7.1. La segmentation est effectuée dans l'espace de couleurs HSV (teinte/saturation/valeur) et nous considérons les canaux de teinte et de valeur. La teinte s'exprime en degrés de 0 à 360 tandis que la valeur vaut entre 0 et 255. La segmentation est effectuée sur l'image 25 de la séquence en prenant comme référence une segmentation de l'image 20. Nous avons choisi une image un peu éloignée de celle de référence pour avoir plus de différences entre les images mais la méthode aurait tout aussi bien marché en cherchant à segmenter l'image suivant celle de référence. Le contour initial est un cercle de rayon 90 pixels au centre de l'image. La spline représentant le contour comporte 64 points d'échantillonnage. Le pas d'évolution est de 0.5, le noyau gaussien de Parzen a une variance de 5. Nous pouvons voir l'évolution du contour sur la figure 7.2. Nous pouvons également suivre l'évolution des *pdf* de l'objet (figure 7.3) et du fond (figure 7.4) au cours du processus de segmentation. Nous voyons que les *pdf* évoluent pour se rapprocher au maximum des *pdf* de référence au fur et à mesure que l'algorithme segmente le visage. Au final, les *pdf* obtenus sont très proches de celles de référence (figure 7.5) et la segmentation du visage est correcte.

Nous testons ensuite notre algorithme sur la séquence **Phone** dans les mêmes conditions. Nous cherchons encore une fois à segmenter le visage du personnage. La segmentation de référence et les deux *pdf* correspondants de l'objet et du fond sont donnés sur la figure 7.6. Le contour initial est un cercle de rayon 60 pixels au centre de l'image, la spline comporte 64 points d'échantillonnage. Le pas d'évolution est de 5, le noyau gaussien de Parzen a une variance de 5. Nous pouvons voir l'évolution de la segmentation sur la figure 7.7 ainsi que l'évolution des *pdf* de l'objet (figure 7.8) et du fond (figure 7.9). A la fin du processus, les *pdf* sont encore très proches de ceux de référence (figure 7.10) et la segmentation du visage est correcte.



(a) Contour de référence

(b) pdf de l'objet (en rouge) et du fond (en vert), présentées en fonction de la teinte et de la valeur(c) pdf de référence de l'objet(d) pdf de référence du fondFIG. 7.1 – Segmentation de référence et pdf de référence de l'objet et du fond

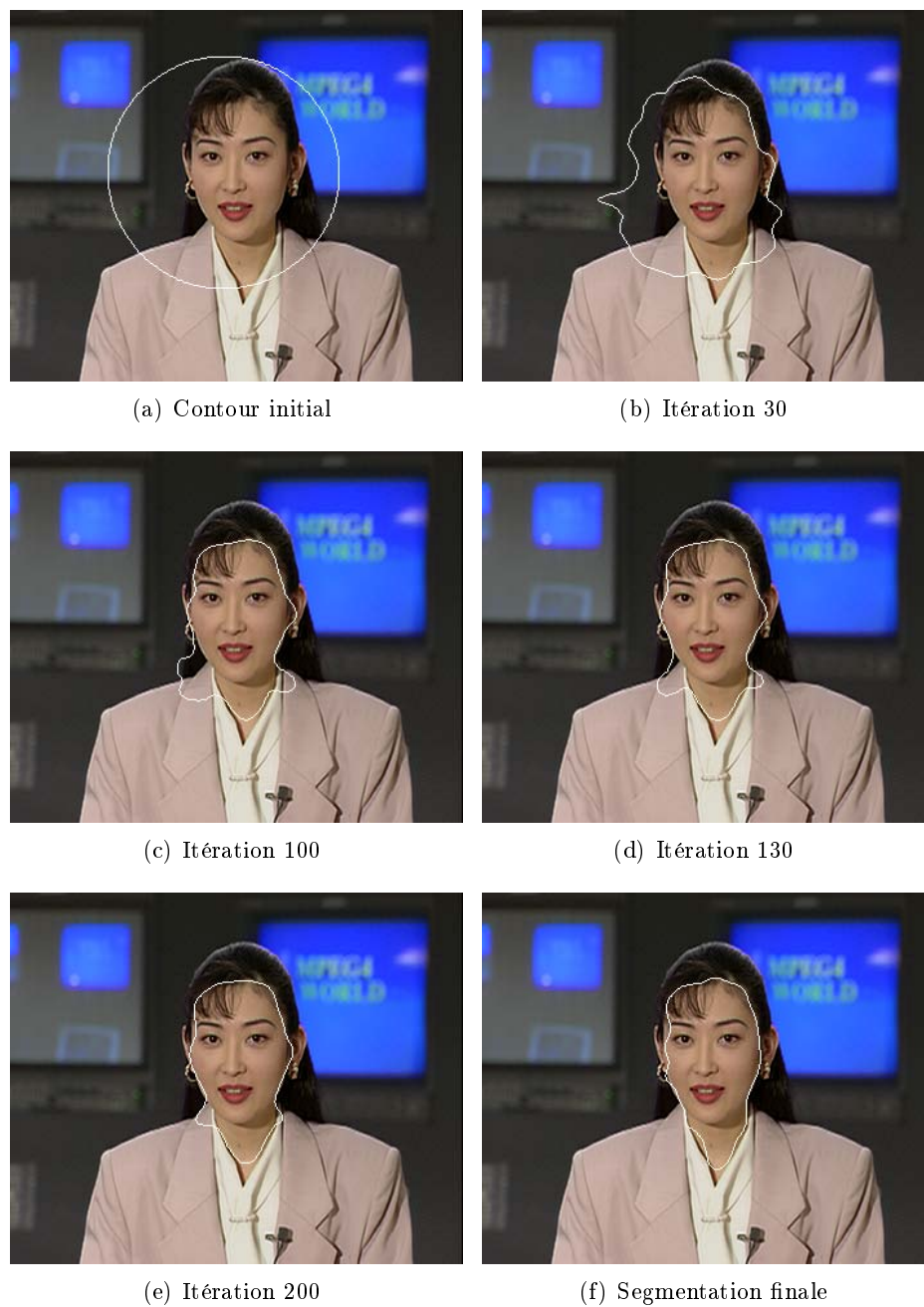
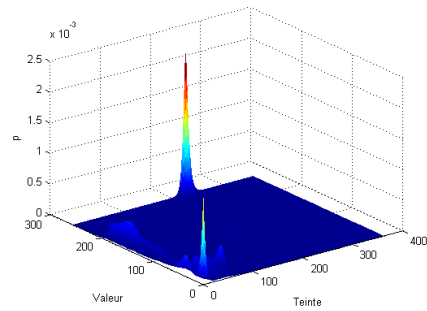


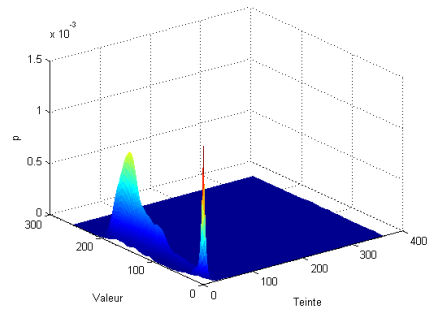
FIG. 7.2 – Evolution de la segmentation



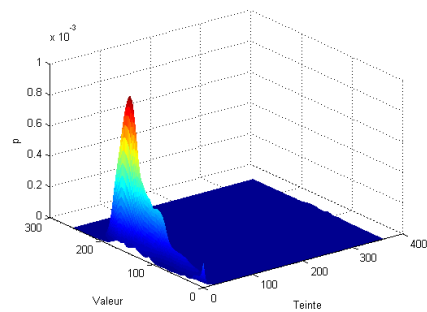
(a) Contour initial

(b) *pdf* initiale de l'objet

(c) Itération 100

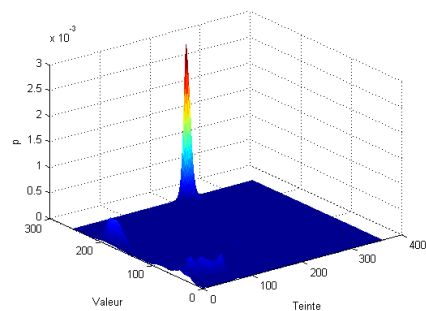
(d) *pdf* de l'objet à l'itération 100

(e) Itération 200

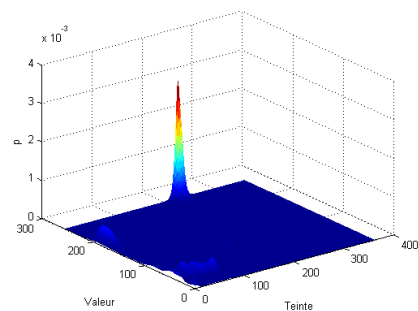
(f) *pdf* de l'objet à l'itération 200FIG. 7.3 – Evolution de la *pdf* de l'objet



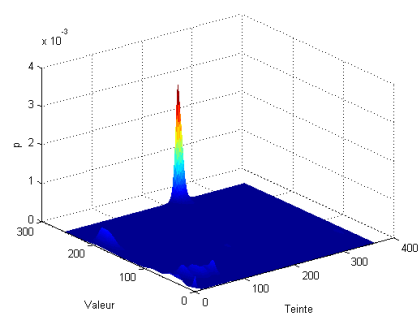
(a) Contour initial

(b) pdf initiale du fond

(c) Itération 100

(d) pdf du fond à l'itération 100

(e) Itération 200

(f) pdf du fond à l'itération 200FIG. 7.4 – Evolution de la pdf du fond

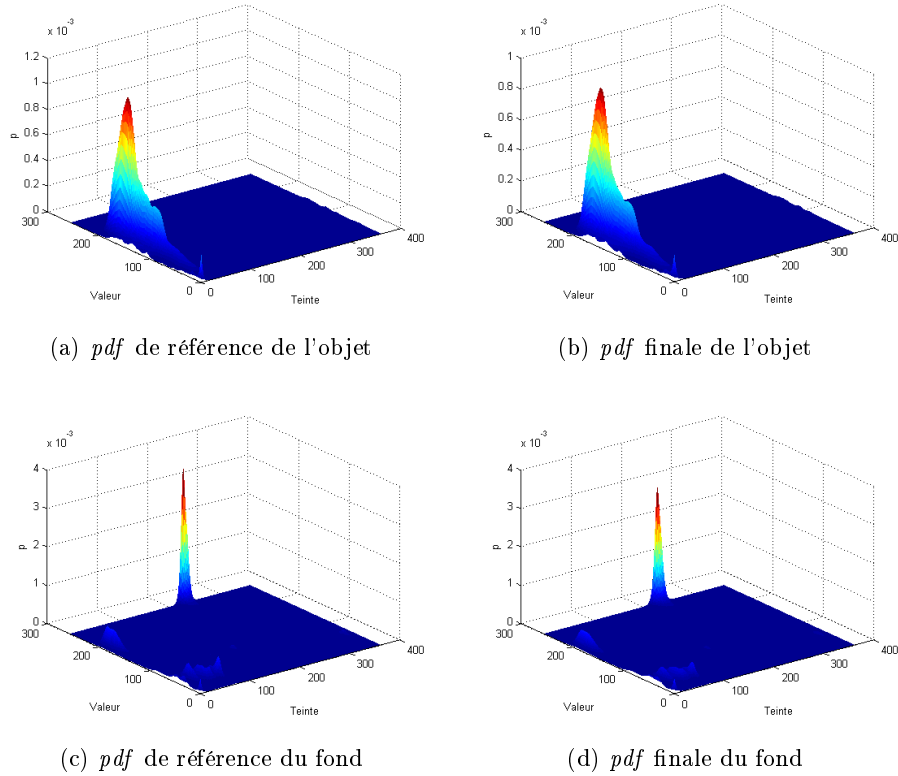


FIG. 7.5 – Comparaison des pdf finales du fond et de l'objet et des pdf de référence

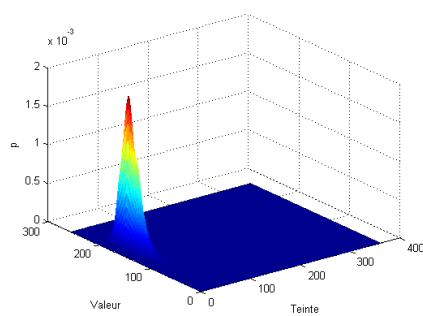
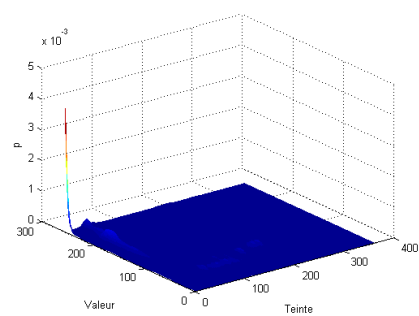
7.3 CONCLUSION

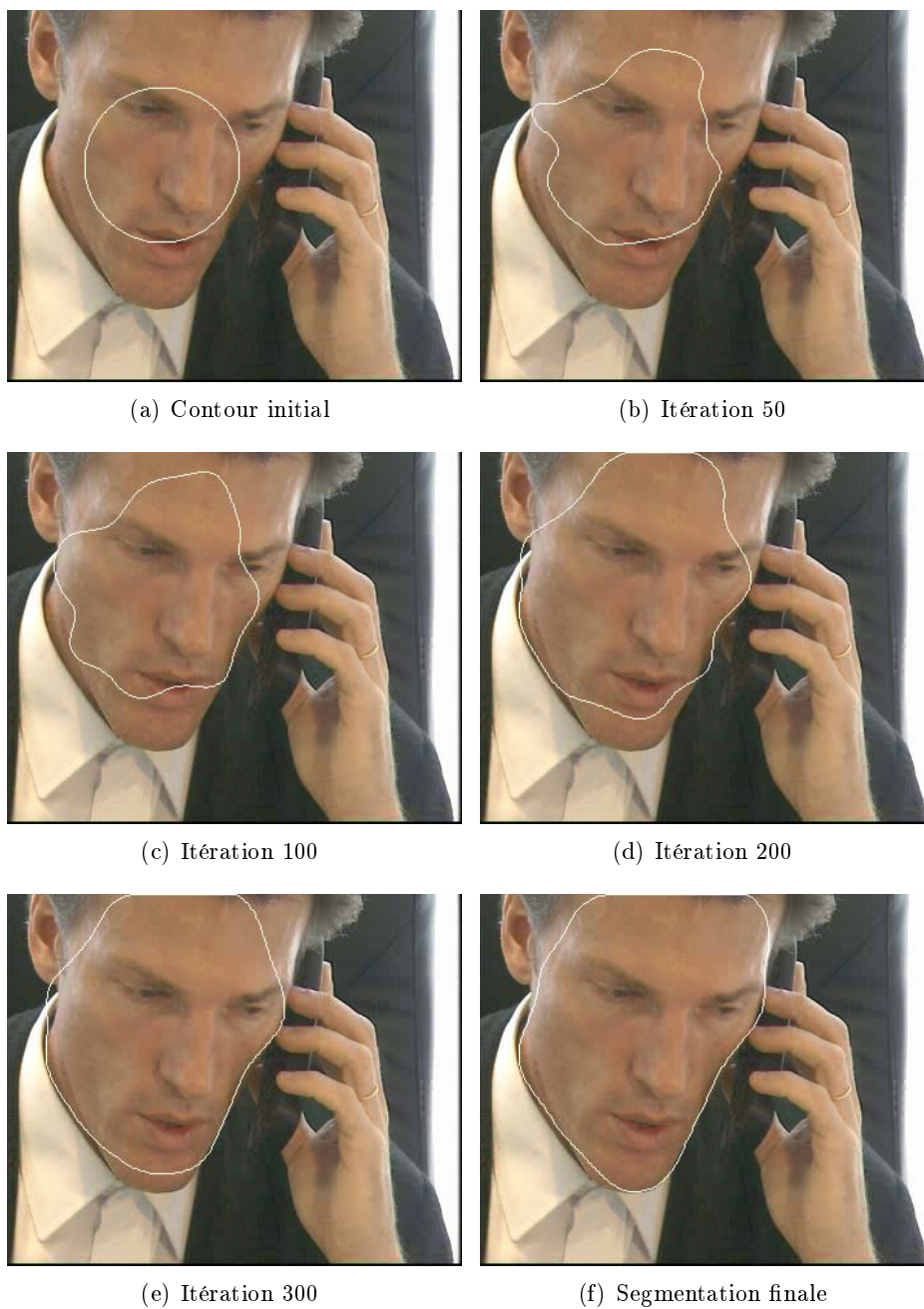
Nous avons utilisé une segmentation de référence afin de segmenter une autre image. Nous avons minimisé une divergence entre densités de probabilité, afin de rendre la densité de probabilité courante la plus proche possible de celle de référence, et de cette façon rendre la région courante (celle comprise à l'intérieur du contour) la plus proche possible de celle de référence en terme de densité de probabilité. Cette méthode a été appliquée à la segmentation de visages dans des séquences vidéo en prenant comme référence une segmentation d'une image de la séquence et en cherchant à segmenter les autres images de la séquence vidéo. Nous avons vu qu'au cours de la segmentation, les densités de probabilité convergent vers les densités de probabilité de référence tandis que le contours actif évolue vers la région concernée.

Il pourrait être intéressant de tester d'autres distances ou divergences, comme la divergence de Jensen-Shannon [AAGLMA⁺00], la distance EMD (*earth mover's distance*) [RTG00], la distance de Kolmogorov [PF00], la distance de Bhattacharyya [GRD04, RAB06, MRT07] ou bien des distances entre histogrammes comme la distance de Wasserstein [CEN07] ou encore des distances entre fonctions de répartition (voir annexe A).



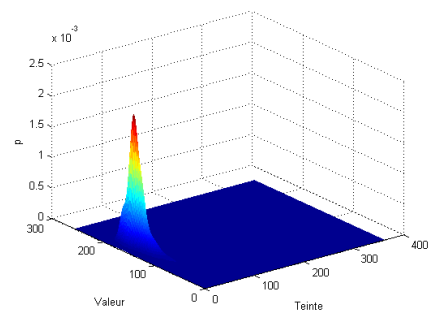
(a) Contour de référence

(b) *pdf* de l'objet (en rouge) et du fond (en vert)(c) *pdf* de référence de l'objet(d) *pdf* de référence du fondFIG. 7.6 – Segmentation de référence et *pdf* de référence de l'objet et du fond

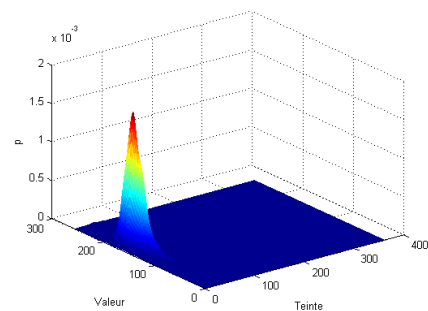
FIG. 7.7 – Evolution de la segmentation sur la séquence **Phone**



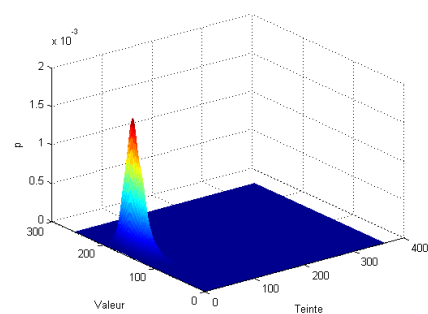
(a) Contour initial

(b) pdf initiale de l'objet

(c) Itération 200

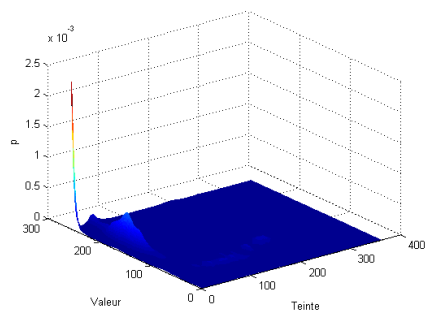
(d) pdf de l'objet à l'itération 200

(e) Itération 900

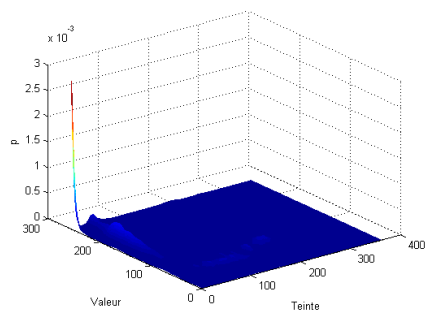
(f) pdf finale de l'objetFIG. 7.8 – Evolution de la pdf de l'objet



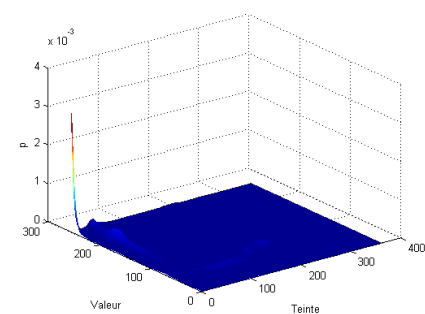
(a) Contour initial

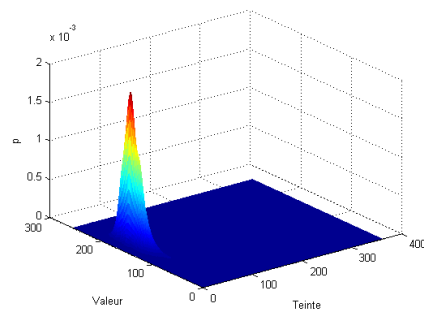
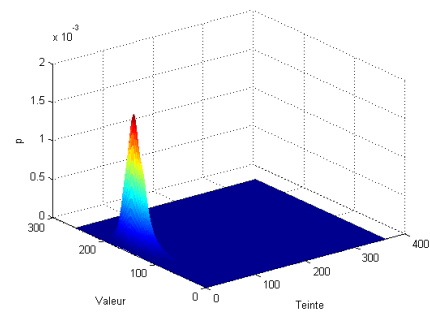
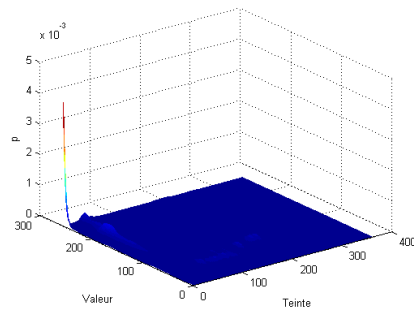
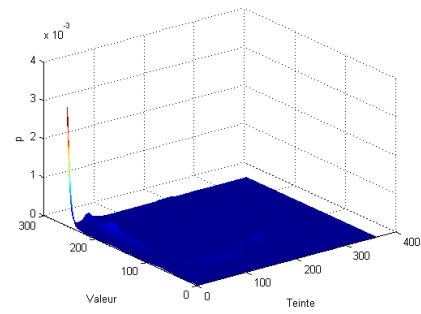
(b) pdf initiale du fond

(c) Itération 200

(d) pdf du fond à l'itération 200

(e) Itération 900

(f) pdf finale du fondFIG. 7.9 – Evolution de la pdf du fond

(a) *pdf* de référence de l'objet(b) *pdf* finale de l'objet(c) *pdf* de référence du fond(d) *pdf* finale du fondFIG. 7.10 – Comparaison des *pdf* finales du fond et de l'objet et des *pdf* de référence

Chapitre 8

Entropie et a priori de distance : exemple avec les watersheds

Dans ce chapitre, nous allons ajouter une contrainte géométrique sous la forme d'un a priori de distance au critère d'entropie. Nous verrons ensuite une application en utilisant une carte de distances venant d'une méthode de segmentation issue de la morphologie mathématique, l'algorithme de la *ligne de partage des eaux*. Cette notion d'a priori géométrique est inspirée des travaux de M. Gastaud [GBA04] et élargie en prenant comme référence un ensemble de contours provenant de l'algorithme de la ligne de partage des eaux (noté *LPE* et appelé *watersheds* en anglais).

Nous rappelons tout d'abord la définition d'une contrainte géométrique, puis nous dérivons ce critère afin de parvenir à l'équation d'évolution du contour actif. Ensuite nous introduisons la notion de distance à une segmentation obtenue par la ligne de partage des eaux. Nous appliquons enfin cette méthode à la segmentation de visages sur des images issues de séquences vidéo.

8.1 DÉFINITION DE LA CONTRAINTE GÉOMÉTRIQUE

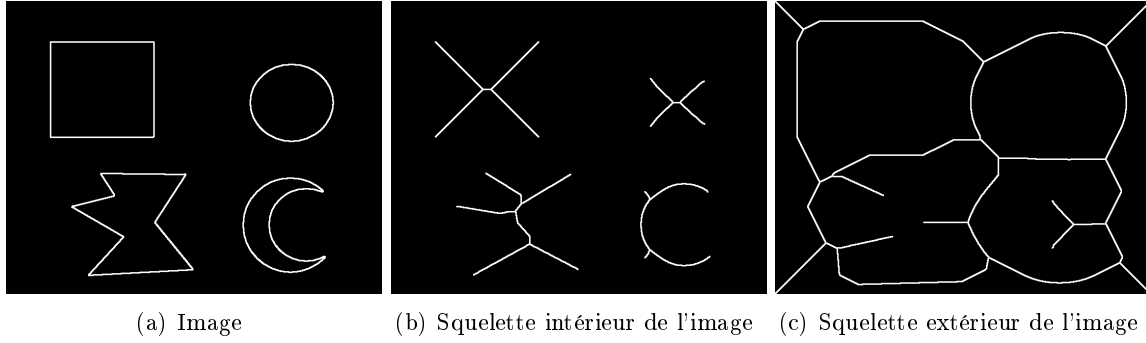
8.1.1 Définition de la distance à un contour

Notons Ω une région de l'image I , délimitée par son contour Γ et soit Γ^{ref} un contour de référence. Notons Ω^{ref} la région comprise à l'intérieur du contour de référence Γ^{ref} . On définit la notion de fonction de distance u d'un point $\mathbf{x}$ au contour Γ^{ref} comme :

$$\begin{aligned} u(\mathbf{x}, \Gamma^{ref}) &= \min_{\mathbf{y} \in \Gamma^{ref}} u(\mathbf{x}, \mathbf{y}) \\ &= \min_{\mathbf{y} \in \Gamma^{ref}} |\mathbf{x} - \mathbf{y}|. \end{aligned} \tag{8.1}$$

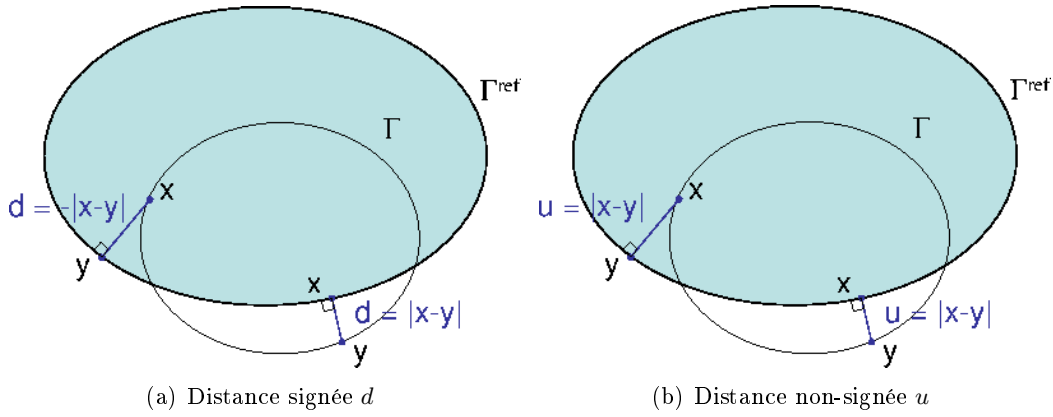
Cette fonction est différentiable presque partout, sauf sur le squelette de Γ^{ref} (voir figure 8.1).

Cette distance peut être définie comme signée ou non-signée (voir figure 8.2). Notons d la distance signée et u la distance non signée. Pour la distance signée, si $\mathbf{x}$ est à l'intérieur de Ω^{ref} , la région délimitée par Γ^{ref} alors sa distance au contour Γ^{ref} est définie négative, s'il est à l'extérieur, sa distance au contour Γ^{ref} est définie positive. Pour une distance



(a) Image (b) Squelette intérieur de l'image (c) Squelette extérieur de l'image

FIG. 8.1 – Image et squelette intérieur et extérieur de cette image.



(a) Distance signée d

(b) Distance non-signée u

FIG. 8.2 – Illustration de la notion de distance signée (à gauche) et de la distance non-signée (à droite) à un contour de référence Γ^{ref} .

non-signée, que $\mathbf{x}$ soit à l'intérieur ou à l'extérieur de la région comprise à l'intérieur de Γ^{ref} , sa distance au contour Γ^{ref} est définie positive.

La fonction distance signée d'un point $\mathbf{x}$ à un contour de référence Γ^{ref} s'écrit donc :

$$d(\mathbf{x}, \Gamma^{ref}) = \begin{cases} -u(\mathbf{x}, \Gamma) = \min_{\mathbf{y} \in \Gamma^{ref}} |\mathbf{x} - \mathbf{y}| & \text{if } \mathbf{x} \in \Omega^{ref} \\ u(\mathbf{x}, \Gamma) = \min_{\mathbf{y} \in \Gamma^{ref}} |\mathbf{x} - \mathbf{y}| & \text{if } \mathbf{x} \notin \Omega^{ref} \end{cases} \quad (8.2)$$

où Ω^{ref} est la région intérieure délimitée par Γ^{ref} .

8.1.2 Critère d'a priori de distance

Une fois la notion de distance définie, nous allons présenter le critère basé sur cette distance. Le but est de contraindre le contour actif en minimisant une distance à un contour de référence. Nous définissons un critère J de la façon suivante :

$$J(\Gamma) = \int_{\Gamma} \varphi(d(\mathbf{x}, \Gamma_{ref})) \, ds \quad (8.3)$$

où s représente l'abscisse curviligne du contour Γ . Le critère J s'écrit donc comme l'intégrale sur le contour Γ d'une fonction φ de la distance d entre un point $\mathbf{x}$ du contour Γ et un contour de référence Γ^{ref} .

Généralement la fonction φ sera choisie de façon à être dérivable, paire et croissante sur $\mathbb{R}^+$. Ainsi, elle permettra de pondérer de la même façon les points à l'intérieur et à l'extérieur du contour et de pénaliser plus fortement les points les plus éloignés du contour.

8.1.3 Dérivation du terme d'a priori géométrique

Afin de pouvoir utiliser le critère (8.3) dans un schéma de contours actifs, nous avons besoin de le dériver pour obtenir l'équation d'évolution du contour actif.

En utilisant les outils de dérivation présentés dans la section 2.3 nous obtenons la dérivée eulérienne dans la direction $\mathbf{V}$ du critère (8.3) :

$$dJ(\Gamma, \mathbf{V}) = \int_{\Gamma} (-\mathbf{N}^{ref} \cdot \mathbf{N} \varphi'(d) - \varphi(d)\kappa)(\mathbf{V} \cdot \mathbf{N}) ds \quad (8.4)$$

où $\mathbf{N}^{ref}$ et $\mathbf{N}$ représentent respectivement les vecteurs normalisés intérieurs du contour de référence Γ^{ref} et du contour actif Γ , κ est la courbure du contour Γ , et $d = d(\mathbf{x}, \Gamma^{ref})$ avec $\mathbf{x} = \Gamma(s)$. (voir figure 8.3 pour la définition des vecteurs normalisés)

La démonstration de cette dérivation est donnée dans [GBA04].

L'équation d'évolution déduite de cette dérivée est donnée par :

$$\frac{\partial \Gamma(\mathbf{s})}{\partial \tau} = (\mathbf{N}^{ref} \cdot \mathbf{N} \varphi'(d) + \varphi(d)\kappa)\mathbf{N} \quad (8.5)$$

où τ est le paramètre d'évolution du contour actif et p paramétrise le contour actif, $d = d(\mathbf{s}, \Gamma^{ref})$, avec $\mathbf{s} = \Gamma(p, \tau)$.

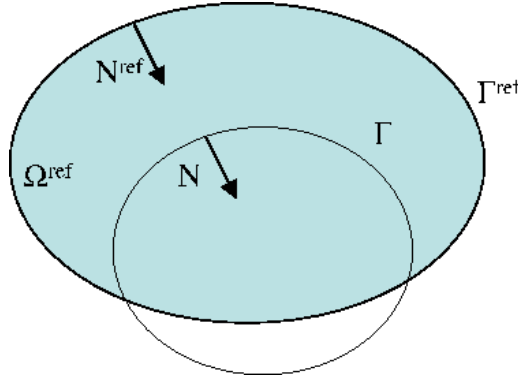


FIG. 8.3 – Définition des vecteurs normalisés intérieurs du contour Γ et du contour Γ^{ref} .

Nous avons ainsi défini la notion de contrainte géométrique en introduisant un critère d'a priori de distance. En dérivant ce critère nous sommes parvenus à l'équation d'évolution pour un tel terme. Nous allons maintenant le combiner à un critère d'entropie afin de rajouter une contrainte géométrique à l'entropie

8.2 ENTROPIE ET A PRIORI DE DISTANCE

Maintenant que nous avons défini un critère d'a priori géométrique, nous allons le combiner avec un critère d'homogénéité présenté chapitre 5, à savoir l'entropie de l'intensité lumineuse. Nous combinons ces deux critères en les sommant de la façon suivante :

$$J(\Omega) = -\frac{1}{|\Omega|} \int_{\Omega} q(I(\mathbf{x}), \Omega) \, d\mathbf{x} + \int_{\Gamma} \varphi(d(\mathbf{x}, \Gamma_{ref})) \, ds \quad (8.6)$$

La dérivée de ce critère s'obtient en combinant les dérivées de chacun de deux termes.

Le choix de la fonction φ doit respecter les contraintes suivantes :

$$\begin{cases} \varphi(0) = 0 \\ \varphi'(0) = 0 \end{cases} \quad (8.7)$$

De plus, l'utilisation de cette fonction permet d'éviter le calcul du gradient de la distance d , qui n'est pas différentiable en un nombre fini de points (sur le squelette de Γ^{ref}).

En rajoutant une contrainte géométrique au terme d'entropie, nous forçons le contour à se rapprocher d'un contour de référence. Nous combinons ainsi les avantages d'un critère d'homogénéité qui permet de segmenter une région qui a des caractéristiques homogènes, et les avantages d'un contour de référence qui force le contour à respecter une certaine contrainte. Ce contour de référence pourrait être une segmentation de référence dans une autre image, une segmentation provenant d'un atlas, une forme de référence ... Nous choisissons ici d'utiliser une segmentation provenant de la ligne de partage des eaux comme contour de référence.

8.3 APPLICATION : DISTANCE PROVENANT DE L'ALGORITHME DE LA LIGNE DE PARTAGE DES EAUX

Nous allons dans cette partie présenter comment nous avons combiné cet a priori géométrique avec une segmentation obtenue par l'algorithme de la ligne de partage des eaux. Nous définissons tout d'abord une distance à partir du résultat de la segmentation par ligne de partage des eaux puis nous utilisons cette distance dans un critère d'a priori géométrique. Enfin nous testons cette nouvelle fonctionnelle pour la segmentation de visages.

8.3.1 Utilisation de l'algorithme de la ligne de partage des eaux pour définir une distance

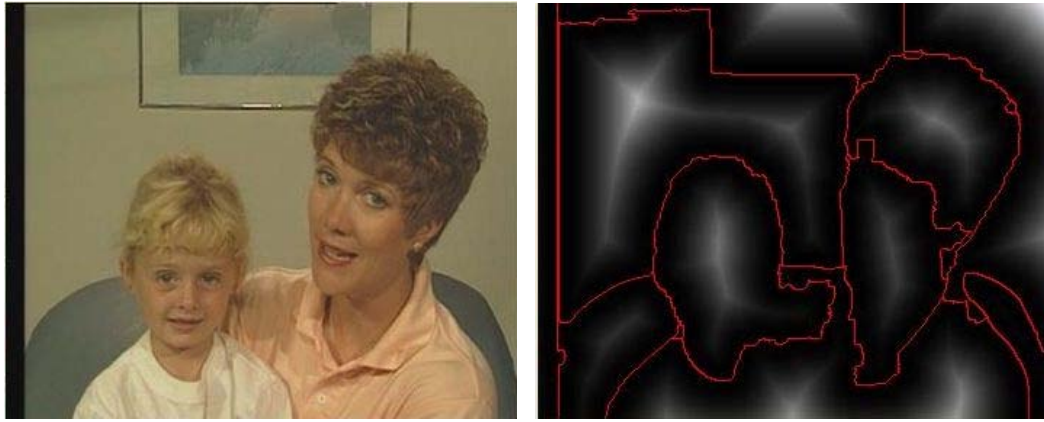
Dans notre approche, nous nous servons de la LPE pour construire une carte de distances qui permettra ensuite de contraindre le contour actif de façon géométrique.

Un rappel de l'algorithme de la ligne de partage des eaux est présenté en annexe B, ainsi que la procédure pour réduire le nombre de régions issues de cet algorithme.

Après avoir obtenu une segmentation avec moins de régions, nous calculons une carte de distances à partir de cette partition. Notons W l'ensemble des lignes de partage des eaux. Nous voulons calculer la distance entre un point $\mathbf{x}$ du contour actif et l'ensemble de courbes W . Dans un ensemble contenant plusieurs régions, nous avons une distance signée pour chaque région. En rassemblant toutes les régions, nous ne gardons plus que l'intérieur de chaque région et nous ne disposons plus que des distances non-signées. Cette distance est donnée par :

$$u(\mathbf{x}, W) = \min_{\mathbf{y} \in W} |\mathbf{x} - \mathbf{y}| \quad (8.8)$$

C'est donc la distance euclidienne d'un point $\mathbf{x}$ à l'ensemble des courbes issues de l'algorithme de la LPE. En pratique, cette distance peut être calculée par le calcul d'une carte de distances. De tels algorithmes renvoient des cartes de distances non signées. Nous pouvons alors appliquer le calcul d'un critère basé sur cette distance, comme présenté dans la section 8.1. La figure 8.4 illustre un exemple de carte de distances sur une image de la séquence **Mother and daughter**. Cette distance est pour un point quelconque la distance à la ligne *watersheds* la plus proche. Elle est toujours positive mais grâce au premier terme de l'équation d'évolution du contour actif, le contour peut aller dans les deux sens, en cherchant à se rapprocher de la ligne *watersheds* la plus proche.



(a) Image de la séquence **Mother and daughter** (b) Carte de distances (plus l'intensité lumineuse est claire, plus le point est éloigné des contours en rouge)

FIG. 8.4 – Carte de distances sur une image de la séquence **Mother and daughter** (en rouge sont représentées les lignes de partage des eaux).

8.3.2 Segmentation contrainte par les watersheds

Une fois définie la distance u , nous pouvons considérer le critère présenté dans la section 8.2 en remplaçant la distance signée d par la distance non-signée u définie en (8.8).

Nous avons donc le critère de segmentation suivant :

$$J(\Omega) = -\frac{1}{|\Omega|} \int_{\Omega} q(I(\mathbf{x}), \Omega) \, d\mathbf{x} + \int_{\Gamma} \varphi(u(\mathbf{x}, W)) \, ds \quad (8.9)$$

Ce critère comporte deux termes, l'un étant l'entropie de l'intensité lumineuse de la région Ω comprise à l'intérieur de la courbe Γ , l'autre terme étant une contrainte de distance à la ligne de partage des eaux la plus proche. Il va permettre de segmenter des zones homogènes en terme de luminosité, tout en forçant le contour à s'approcher des zones à fort contour, le terme de contrainte de distance attirant le contour vers les endroits où le gradient est fort.

Le choix de la fonction φ doit respecter quelques contraintes, comme nous l'avons vu précédemment. La fonction doit être dérivable, paire sur $\mathbb{R}^+$ et doit s'annuler ainsi que sa dérivée en 0. Pour les expérimentations, nous avons choisi la fonction carré : $\varphi(t) = t^2$.

Le critère (8.9) est dérivable presque partout, sauf sur le squelette de la carte de distances, *i.e.* en un nombre fini de points.

Nous obtenons alors l'équation d'évolution suivante pour le contour actif :

$$\begin{aligned} \frac{\partial \Gamma(p, \tau)}{\partial \tau} &= (\mathbf{N}^{ref} \cdot \mathbf{N} \, 2u + \kappa u^2) \mathbf{N} \\ \Gamma(0, \tau) &= \Gamma_0(\tau) \end{aligned} \quad (8.10)$$

avec $u = u(\mathbf{x}, W)$ et $\mathbf{x} = \Gamma(p, \tau)$.

8.3.3 Résultats expérimentaux

Nous testons la segmentation d'images issues de séquences vidéo en utilisant le critère (8.9). Nous comparons les résultats obtenus avec et sans la contrainte de distance pour la segmentation de zones homogènes.

Le premier test est effectué sur une image de la séquence **Foreman**, nous cherchons à segmenter le visage. Nous effectuons au préalable une segmentation par LPE pour créer la carte de distances qui est utilisée comme critère par l'algorithme de contour actif. Nous choisissons de créer la carte de distances avec une segmentation de l'image en 30 régions. Le nombre de régions est arbitraire et dépend de l'image considérée. L'idée est de fusionner les régions obtenues après la ligne de partage des eaux en un nombre qui représente les zones principales de l'image, *i.e.* de fusionner au maximum les zones de même couleur par exemple. La figure 8.5 présente l'image que l'on cherche à segmenter et la carte de distances calculée à partir de la partition de cette image en 30 régions.

La figure 8.6 présente l'évolution du contour actif avec et sans la contrainte de distance provenant de l'algorithme de la LPE. L'initialisation est un cercle de rayon 80 pixels, centré sur l'image. La spline qui comporte 128 points d'échantillonnage. Le pas d'évolution est de 0.5 et la fenêtre de Parzen est une gaussienne de moyenne nulle et de variance 5. Nous voyons que sans la contrainte géométrique, le contour déborde du visage et inclut une

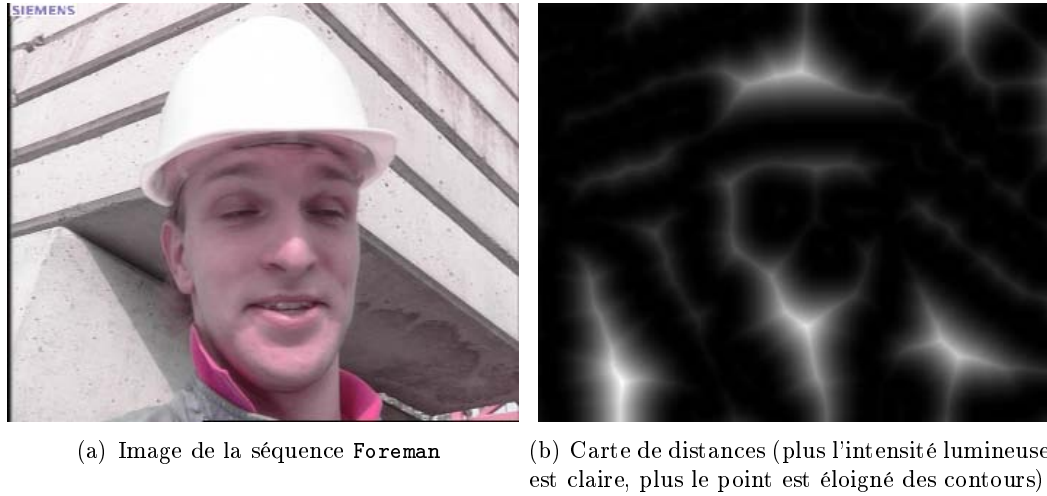


FIG. 8.5 – Image issue de la séquence **Foreman** et carte de distances résultant d'une partition de l'image en 30 régions par l'algorithme des watersheds.

partie du fond tandis qu'avec la contrainte de distance, le contour segmente correctement le visage du personnage.

La figure 8.7 présente quant à elle l'évolution du contour actif sur une image de la séquence **Mother and daughter**. La carte de distances est calculée à partir d'une partition de l'image en 11 régions. L'initialisation est un cercle de rayon 85 pixels, positionné en (160, 230). La spline représentant le contour comporte 128 points d'échantillonnage. Le pas d'évolution est de 1.5 et la fenêtre de Parzen est une gaussienne de moyenne nulle et de variance 5. Nous pouvons remarquer que sans la contrainte géométrique, le contour actif englobe aussi le visage de la petite fille et qu'il segmente peu précisément le visage de la mère. Avec la contrainte géométrique, nous obtenons une segmentation précise du visage de la mère.

8.4 CONCLUSION

Nous avons présenté un critère de segmentation basé sur un *a priori* de distance. Nous avons combiné un terme d'entropie pour segmenter une région homogène, et un terme de distance afin de se rapprocher des contours de l'image. Ce terme de distance est ici issu d'une segmentation préalable obtenue par la méthode de la ligne de partage des eaux. Cette méthode donne une segmentation basée sur le gradient de l'image et comporte beaucoup de régions. En combinant l'entropie et cette distance à la segmentation obtenue par la ligne de partage des eaux, nous améliorons la qualité de la segmentation par rapport à celle obtenue avec un terme d'entropie seul.

Il pourrait être intéressant de combiner l'entropie à d'autres contraintes géométriques, comme par exemple des formes de référence afin de segmenter des objets ayant une forme

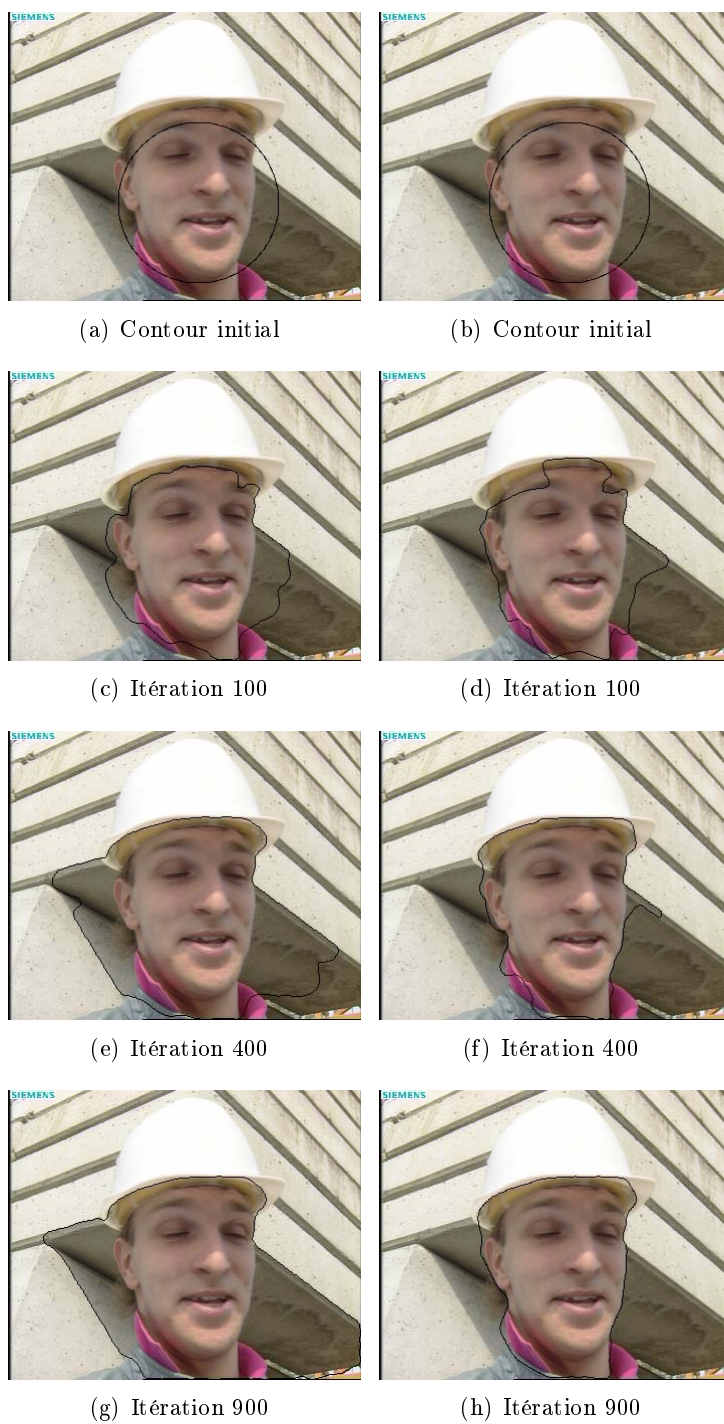


FIG. 8.6 – Evolution du contour actif sans la contrainte géométrique (colonne de gauche) et avec la contrainte géométrique (colonne de droite).

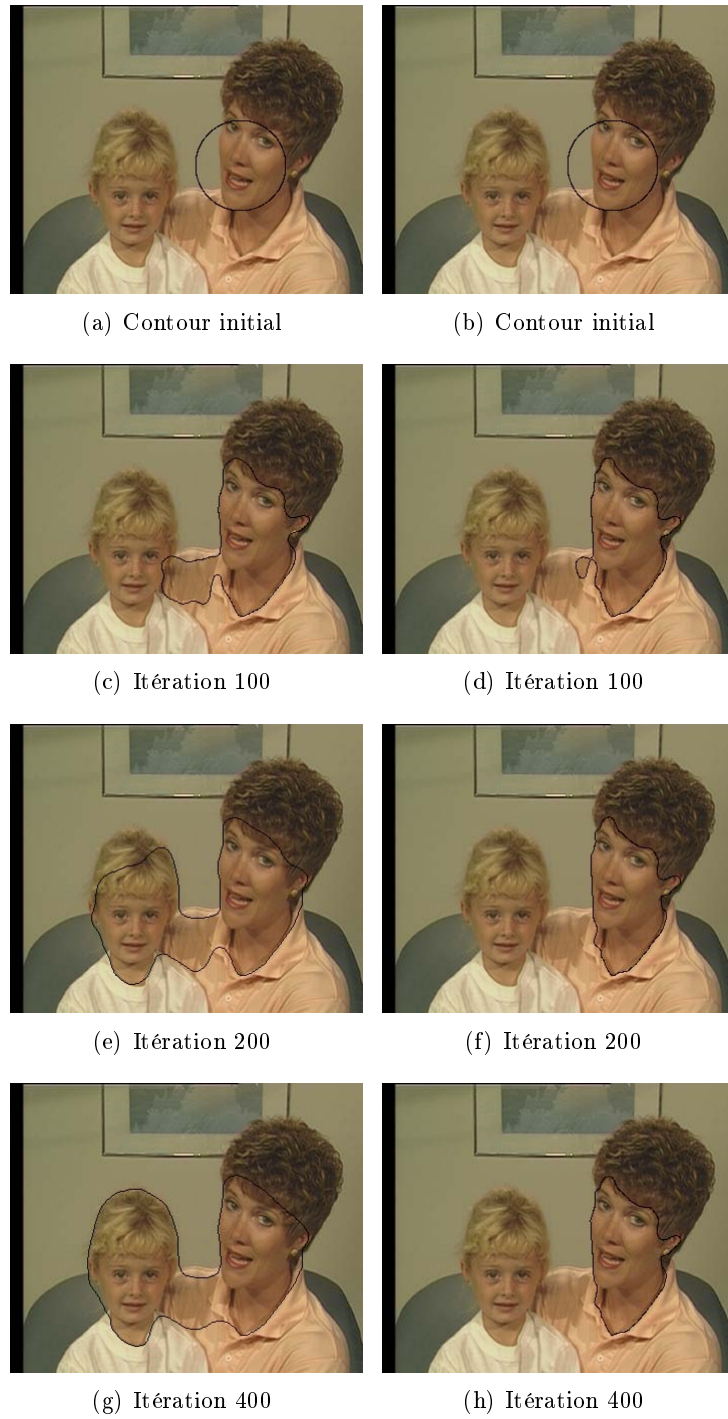


FIG. 8.7 – Evolution du contour actif sans la contrainte géométrique (colonne de gauche) et avec la contrainte géométrique (colonne de droite).

particulière. Le critère d'entropie garantirait une certaine homogénéité de l'objet tandis que la contrainte géométrique assurerait une forme prédéfinie.

Troisième partie

Segmentation de séquences vidéo

Chapitre 9

Introduction à la segmentation de vidéos

La segmentation d'objets dans des séquences vidéo est une tâche difficile. Les objets en mouvement ne sont pas faciles à segmenter en raison des occultations, des changements de luminance, du fait que ces objets peuvent être articulés et des mouvements de caméra. Ses applications peuvent être aussi bien la télésurveillance, l'imagerie médicale, la météorologie ou l'indexation. De plus, dans les normes de codage comme MPEG-4 et MPEG-7, le découpage de la vidéo se fait en objets vidéo et non plus en blocs mais cela n'est toujours pas mis en pratique donc la recherche sur la segmentation et le suivi d'objets en mouvement a pris de l'importance.

Nous pouvons séparer les méthodes de détection de mouvement des méthodes d'estimation du mouvement. Une revue de ces méthodes peut être trouvée dans [Zha01].

9.1 DÉTECTION DE MOUVEMENT

Dans les méthodes de détection de mouvement, le mouvement des objets n'a pas besoin d'être calculé, les objets sont caractérisés par le fait que ce sont les seules parties de la vidéo qui ont un mouvement important. Il s'agit d'abord de distinguer les séquences à caméra fixe et les séquences à caméra mobile.

Dans les séquences à caméra fixe, seuls les objets en mouvement se déplacent, le fond de la séquence restant relativement statique. Il peut y avoir des éléments du fond qui ont un mouvement, comme par exemple des nuages dans le ciel ou bien des branches d'arbre qui se déplacent avec le vent mais ce ne sont pas les mouvements dominants dans la scène et la plupart des pixels du fond restent statiques.

Les méthodes les plus simples de détection de mouvement consistent à calculer l'image différence entre deux images successives de la séquence vidéo. Les pixels qui ont le plus fort mouvement sont considérés comme faisant partie de l'objet en mouvement. C'est cette approche dans un contexte statistique qui est présenté dans [PD00].

On peut aussi faire la différence entre une image et une image de référence. Par exemple, dans une application de télésurveillance, on peut disposer d'images de référence (les images quand il n'y a personne dans la scène) et considérer les éléments différents de l'image

courante par rapport à l'image de référence comme les objets d'intérêt. Cette image de référence peut être celle du fond [PD99].

Il est aussi possible de restaurer l'image tout en segmentant les objets en mouvement [KDA99].

Dans les séquences à caméra mobile, il est nécessaire d'estimer le mouvement de la caméra afin de le compenser pour pouvoir segmenter les objets en mouvement. Le mouvement dominant de la scène appliqué à tous les pixels est considéré comme étant le mouvement de la caméra et les pixels ne suivant pas ce mouvement global forment les objets en mouvement. Le principe de la compensation de caméra est expliqué dans l'annexe C. Il faut donc compenser le mouvement de la caméra afin de se replacer dans un cadre de caméra fixe. L'idée est d'utiliser une estimation de caméra la plus robuste possible [CBFAB97, OB95]. Wu et Kittler [WK93] utilisent un modèle de mouvement affine pour représenter le mouvement de la caméra et détectent les objets en mouvement par un seuillage. De son côté, Salembier [SM99] propose une approche morphologique pour la segmentation d'objets en mouvement.

Certains ont utilisés des méthodes d'estimation du fond de la séquence vidéo (la partie statique) afin de segmenter les objets en mouvement. En caméra fixe, Jehan-Besson *et al.* [JbBA02] ont proposé d'estimer le fond de la séquence à partir de n images de la séquence. Le fond peut aussi être calculé par une approche statistique en modélisant l'image de différences comme une mixture de laplaciennes [ST01]. En caméra mobile, on appelle ce fond une mosaïque [IAB⁺96, GB02].

9.2 ESTIMATION DU MOUVEMENT

D'autres méthodes utilisent l'estimation d'un champ dense de mouvement, c'est-à-dire l'estimation d'un vecteur mouvement par pixel de l'image. Ce sont les méthodes basées sur le *flot optique*, qui repose sur le fait qu'il y a invariance d'illumination d'un objet au cours du temps. Dans ce type d'estimation, il y a beaucoup d'inconnues, à savoir deux fois le nombre de pixels et il y a une seule contrainte scalaire par pixel. C'est donc un problème mal posé qui a besoin d'être régularisé. D'un côté certaines méthodes conservent un champ dense de vecteurs mais imposent des conditions de lissage de façon globale sur le champ de vecteurs. C'est le cas notamment de [HS81, AWS99, WS01]. D'un autre côté, des méthodes locales contraignent le champ de vecteurs à suivre une hypothèse paramétrique de mouvement (par exemple un mouvement de translation, affine, ou une homographie) sur toute l'image ou sur certaines parties de celle-ci [LK81, OB95, WK93]. Une autre méthode combine ces deux approches en utilisant un champ dense de vecteurs combiné à la robustesse au bruit des méthodes locales [BWS02].

Le flot optique peut servir d'étape préalable à une segmentation [PBMR00, RD04, RD05, HJBD⁺06]. Le flot optique est ainsi estimé de façon globale sur l'image considérée puis cette image est segmentée par contours actifs.

L'estimation du mouvement de l'objet et l'estimation de ses frontières peuvent aussi être calculées en même temps [BF93, MP98b, Far99, MK99, SY02, CS03, DGBA05, CS05, VML06]. De plus le suivi d'objets mobiles au cours de la séquence peut amener l'étude de critères spatio-temporels pour tenir compte des propriétés spatiales de l'objet considéré [MEFM02, BRDW03].

9.3 PRÉSENTATION DES TRAVAUX SUR LA SEGMENTATION DE SÉQUENCES VIDÉO

Dans cette partie nous présentons deux méthodes de segmentation d'objets en mouvement dans des séquences vidéo. Tout d'abord dans le chapitre 10, nous utilisons un champ de vecteurs dense calculé au préalable pour minimiser l'entropie de ses composantes. Nous segmentons de cette façon des objets en mouvement mais la précision de cette segmentation repose sur la qualité du flot optique obtenu et de ce fait, certaines portions des objets peuvent être laissées de côté.

Dans le chapitre 11, nous proposons une autre méthode de segmentation d'objets mobiles. Cette fois-ci nous ne calculons plus de champ dense mais nous cherchons à estimer conjointement le mouvement des objets et leur segmentation. Pour ce faire nous minimisons une fonctionnelle spatio-temporelle prenant en compte le mouvement et l'intensité des objets d'intérêt. Cette méthode a l'avantage de combiner le spatial et le temporel et permet de segmenter correctement des objets en mouvement même si des séquences très difficiles peuvent mettre en défaut le modèle proposé.

Segmentation de données vectorielles : entropie du flot optique

Dans ce chapitre nous proposons de segmenter des objets en mouvement dans des séquences vidéo à caméra fixe ou mobile. Nous considérons le flot optique des séquences vidéo à segmenter, c'est-à-dire que nous calculons un vecteur mouvement par pixel pour chaque paire d'images consécutives. Nous faisons face à un champ dense de vecteurs mouvement. Les caractéristiques que nous considérons sont donc les coordonnées du vecteur mouvement : $\mathbf{f}(\mathbf{x}) = [u, v]^T$. Ce n'est pas uniquement la longueur du vecteur mouvement que nous allons prendre en compte, mais aussi sa direction.

Nous présentons tout d'abord le critère de segmentation que nous considérons, à savoir l'entropie des composantes des vecteurs mouvement. Puis nous testons ce critère sur des champs de mouvement synthétiques afin de valider la méthode. Enfin, nous appliquons ce critère à des séquences réelles dont le mouvement est estimé par flot optique.

10.1 L'ENTROPIE POUR SEGMENTER LE MOUVEMENT

Nous cherchons à segmenter des objets mobiles dans des séquences vidéo. Pour ce faire, nous calculons le flot optique de la séquence vidéo puis nous cherchons à minimiser l'entropie des vecteurs mouvement obtenus avec le flot optique. Nous utilisons donc l'entropie de données vectorielles comme dans le chapitre 6 mais cette fois les données vectorielles sont des vecteurs mouvement.

10.1.1 Entropie

Nous considérons donc le critère de segmentation suivant :

$$H(\Omega) = -\frac{1}{|\Omega|} \int_{\Omega} \log p(\mathbf{f}(\mathbf{x}), \Omega) d\mathbf{x} \quad (10.1)$$

avec $\mathbf{f}$ le vecteur mouvement en chaque point à savoir $\mathbf{f} = [u, v]^T$.

Le critère à minimiser est l'entropie des vecteurs mouvement en chaque point de la région Ω . Nous cherchons des régions où les vecteurs mouvement sont les plus homogènes

possibles. De ce fait nous allons séparer les objets qui ont un mouvement différent.

Pour minimiser le critère (10.1), nous avons besoin de le dériver. Nous rappelons l'expression de la dérivée eulérienne dans la direction $\mathbf{V}$ d'un tel critère :

$$dH(\Omega, \mathbf{V}) = - \int_{\Gamma} \frac{1}{|\Omega|} (A(s, \Omega) + \varphi(p(\mathbf{f}(s), \Omega))) (\mathbf{V} \cdot \mathbf{N}) ds . \quad (10.2)$$

avec

$$A(s, \Omega) = - \left(H(\Omega) + 1 + \frac{1}{|\Omega|} \int_{\Omega} \frac{K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(s))}{p(\mathbf{f}(\mathbf{x}), \Omega)} d\mathbf{x} \right) .$$

et K_h est le noyau gaussien de moyenne nulle et de variance h .

De cette dérivée nous en déduisons l'équation d'évolution du contour actif :

$$\frac{\partial \Gamma}{\partial \tau}(\mathbf{s}) = \frac{1}{|\Omega|} (A(\mathbf{s}, \Omega) + -\log(p(\mathbf{f}(\mathbf{s}), \Omega))) \mathbf{N} \quad (10.3)$$

avec $\mathbf{s} = \Gamma(p, \tau)$, $\mathbf{N}$ la normale interne au contour actif et :

$$A(\mathbf{s}, \Omega) = - \left(H(\Omega) + 1 + \frac{1}{|\Omega|} \int_{\Omega} \frac{K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{s}))}{p(\mathbf{f}(\mathbf{x}), \Omega)} d\mathbf{x} \right)$$

10.1.2 Entropie conditionnelle

Dans les expérimentations, nous minimisons un critère correspondant à l'entropie de l'image sachant que l'on se situe à l'intérieur ou à l'extérieur de la courbe. Cela représente donc l'entropie conditionnellement à un label d'appartenance à la région. La fonctionnelle est alors :

$$J(\Gamma) = |\Omega_{in}| H(\Omega_{in}) + |\Omega_{out}| H(\Omega_{out}) \quad (10.4)$$

avec Ω_{in} la région comprise à l'intérieur de la courbe et Ω_{out} la région comprise à l'extérieur.

La dérivée eulérienne de cette fonctionnelle dans la direction $\mathbf{V}$ est donnée par :

$$\begin{aligned} dJ(\Gamma, \mathbf{V}) = & \int_{\Gamma} \left(\log(p(\mathbf{f}(s), \Omega_{in})) + \frac{1}{|\Omega_{in}|} \int_{\Omega_{in}} \frac{K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(s))}{p(\mathbf{f}(\mathbf{x}), \Omega_{in})} d\mathbf{x} - 1 \right. \\ & \left. - (\log(p(\mathbf{f}(s), \Omega_{out})) + \frac{1}{|\Omega_{out}|} \int_{\Omega_{out}} \frac{K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(s))}{p(\mathbf{f}(\mathbf{x}), \Omega_{out})} d\mathbf{x} - 1) (\mathbf{V} \cdot \mathbf{N}) ds \right) \end{aligned} \quad (10.5)$$

Nous pouvons alors en déduire l'équation d'évolution du contour actif :

$$\begin{aligned} \frac{\partial \Gamma}{\partial \tau}(\mathbf{s}) = & \left(\log \frac{p(\mathbf{f}(\mathbf{s}), \Omega_{out})}{p(\mathbf{f}(\mathbf{s}), \Omega_{in})} - \frac{1}{|\Omega_{in}|} \int_{\Omega_{in}} \frac{K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{s}))}{p(\mathbf{f}(\mathbf{x}), \Omega_{in})} d\mathbf{x} \right. \\ & \left. + \frac{1}{|\Omega_{out}|} \int_{\Omega_{out}} \frac{K_h(\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{s}))}{p(\mathbf{f}(\mathbf{x}), \Omega_{out})} d\mathbf{x} \right) \mathbf{N} \end{aligned} \quad (10.6)$$

Dans cette équation, le premier terme est prépondérant. Si la densité de probabilités $p(\mathbf{f}(\mathbf{s}), \Omega_{out})$ est supérieure à $p(\mathbf{f}(\mathbf{s}), \Omega_{in})$ la courbe va avancer vers l'intérieur pour exclure

le pixel considéré. Ainsi si le mouvement calculé au point $\mathbf{s}$ a une plus forte probabilité d'appartenir au fond plutôt qu'à l'objet, l'algorithme va faire évoluer la courbe de façon à ne pas inclure ce point dans la région comprise à l'intérieure de la courbe.

Nous allons utiliser cette équation pour segmenter des objets en mouvement, tout d'abord en considérant des mouvements synthétiques puis en considérant des séquences d'images réelles et en calculant le mouvement par flot optique.

10.2 RÉSULTATS SUR DES IMAGES SYNTHÉTIQUES

Dans cette section, nous présentons des résultats de segmentation sur des champs de vecteurs mouvement synthétiques en utilisant un descripteur basé sur l'entropie conditionnelle. Nous testons tout d'abord le critère sur des mouvements synthétiques afin de valider la méthode avant de la confronter à des séquences réelles.

Pour ne pas prendre en compte que la longueur des vecteurs mouvement mais aussi leur direction, les coordonnées $2D$ (u, v) des vecteurs sont utilisées.

Les figures suivantes représentent seulement un vecteur sur 10, et pour plus de lisibilité, leur taille est multipliée par 8. Le point de départ de chaque vecteur est le pixel à partir duquel le vecteur est calculé.

Le premier exemple montre un rectangle se déplaçant vers la droite sur un fond se déplaçant vers la gauche (figure 10.1(a)). Les vecteurs du rectangle et du fond ont la même taille donc la longueur du vecteur ne peut pas nous donner d'information discriminante pour la segmentation dans ce cas. Mais comme nous prenons en compte la direction du vecteur, la segmentation du rectangle peut se faire correctement. La figure 10.1(c) montre la segmentation finale de cet exemple.

Dans le second exemple, trois rectangles de taille différente ayant un mouvement différent, sont situés sur un fond qui se déplace (figure 10.2(a)). La figure 10.2(e) montre le résultat de la segmentation. Un petit rectangle débordant sur les trois différents champs de vecteurs est choisi comme initialisation donc il va devoir s'étendre à certains endroits et rétrécir à d'autres. De plus, il va y avoir un changement de topologie, c'est-à-dire qu'il va devoir se scinder en plusieurs morceaux. L'évolution des histogrammes est présentée sur la figure 10.3. La colonne de gauche montre l'évolution de l'histogramme du fond, tandis que celle de droite présente celle de l'histogramme des objets. De façon claire, le pic en $(0, 1)$ correspond au mouvement du fond qui disparaît graduellement dans l'histogramme de l'objet, tandis que dans l'histogramme du fond il reste le seul pic. Cela montre que notre approche réussit à segmenter plusieurs objets ayant un mouvement différent sur un fond dont le mouvement est homogène.

Le dernier exemple synthétique montre un disque qui s'élargit sur un fond qui se déplace vers la droite (figure 10.4(a)). Ici les vecteurs mouvement de l'objet vont dans presque toutes les directions. La longueur de ces vecteurs n'est pas non plus constante à l'intérieur de l'objet. Malgré cela, notre méthode est capable de segmenter correctement le disque du

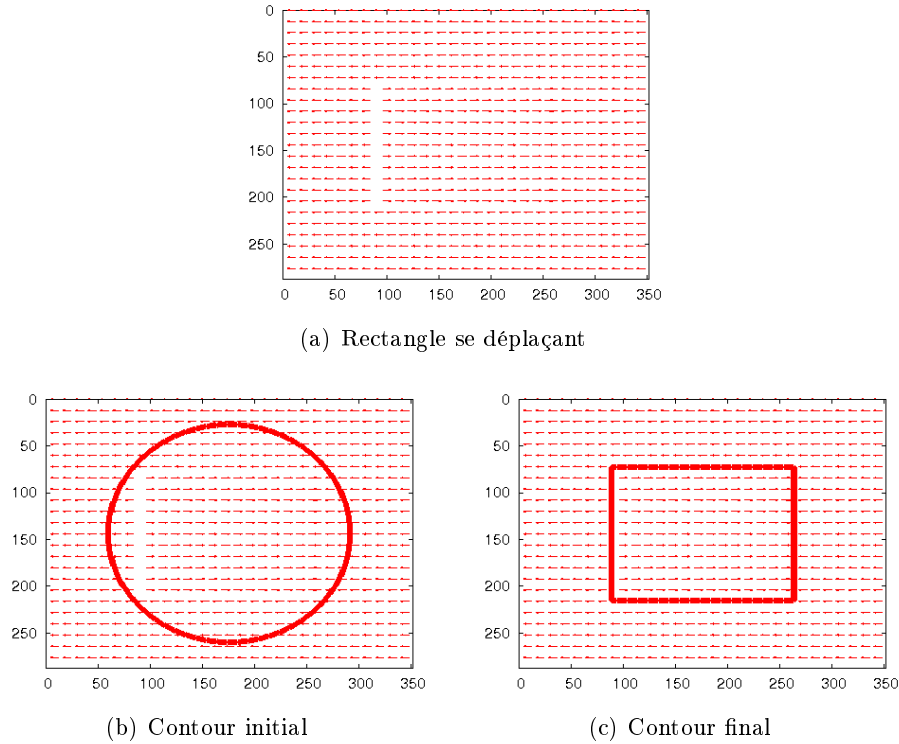


FIG. 10.1 – Rectangle se déplaçant : Segmentation

fond parce que le mouvement du fond est homogène. La figure 10.4 illustre ce résultat.

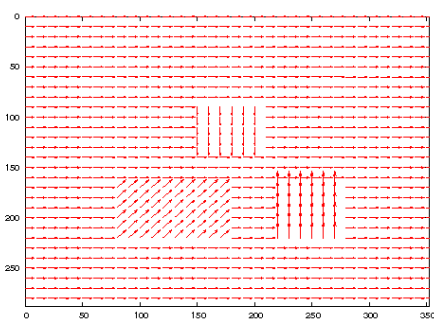
10.3 RÉSULTATS SUR DES SÉQUENCES RÉELLES

Dans la partie précédente, nous avons étudié le cas de champs de vecteurs mouvement synthétiques et nous avons pu observer le résultat de la segmentation de champs de mouvement simples. Pour segmenter des séquences réelles, il nous faut estimer le mouvement entre deux images successives de la séquence. Pour cela, nous calculons le flot optique entre deux paires d'images. Notons que nous pourrions utiliser n'importe quel méthode d'estimation de mouvement comme entrée pour notre algorithme de segmentation.

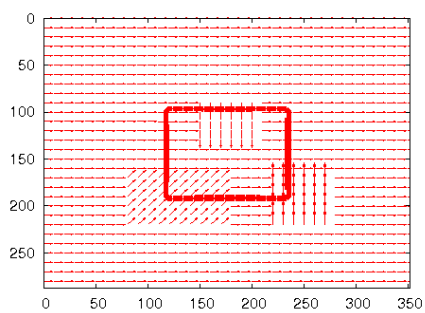
Nous allons donc d'abord présenter l'estimation de mouvement par flot optique puis nous allons minimiser l'entropie des composantes des vecteurs mouvement obtenus. Nous allons appliquer cette méthode à la segmentation d'objets en mouvement dans des séquences à caméra fixe et à caméra mobile.

10.3.1 Estimation du mouvement par flot optique

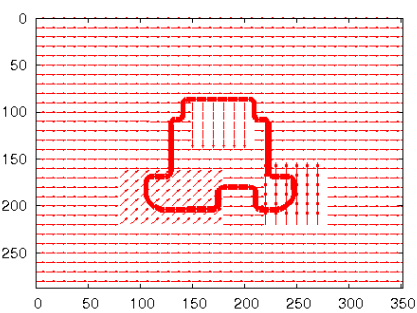
Dans ce qui suit, une séquence d'images est notée $I(x, y, t)$ où (x, y) représente la localisation dans le domaine de l'image Ω et t le temps. Une façon d'estimer le mouvement dans une séquence d'images est de calculer le flot optique, c'est-à-dire de calculer pour



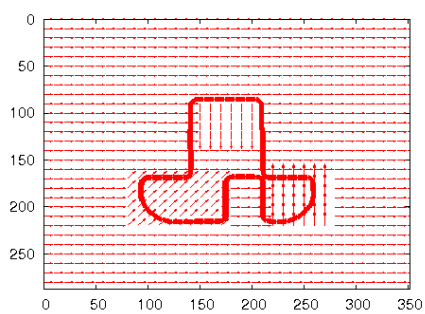
(a) Rectangles se déplaçant



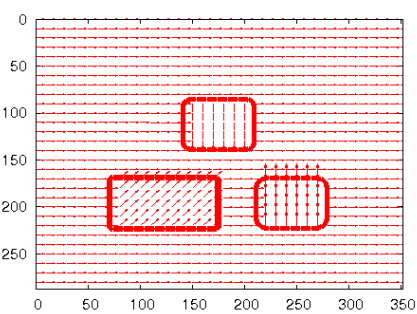
(b) Contour initial



(c) Itération 40

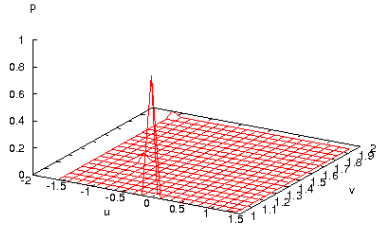


(d) Itération 80

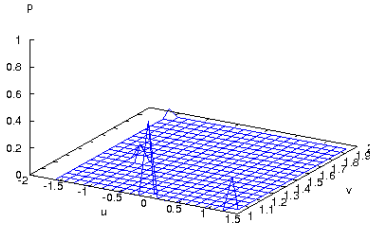


(e) Contour final

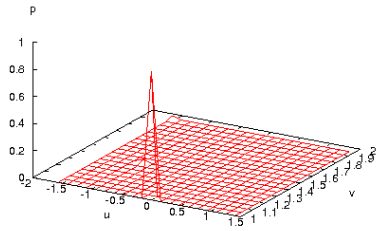
FIG. 10.2 – Rectangles se déplaçant : Segmentation



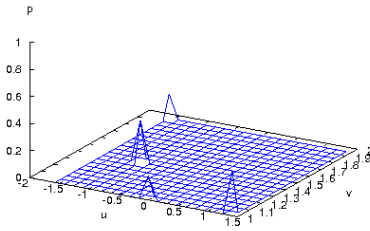
(a) Histogramme initial du fond



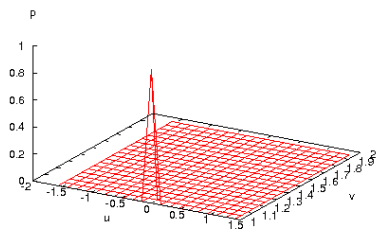
(b) Histogramme initial de l'objet



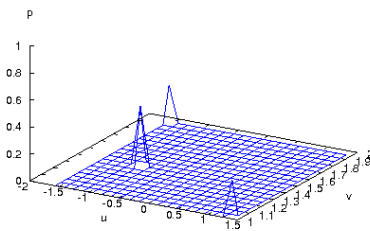
(c) Histogramme initial du fond à l'itération 80



(d) Histogramme initial de l'objet à l'itération 80



(e) Histogramme final du fond



(f) Histogramme final de l'objet

FIG. 10.3 – Rectangles se déplaçant : Evolution des histogrammes

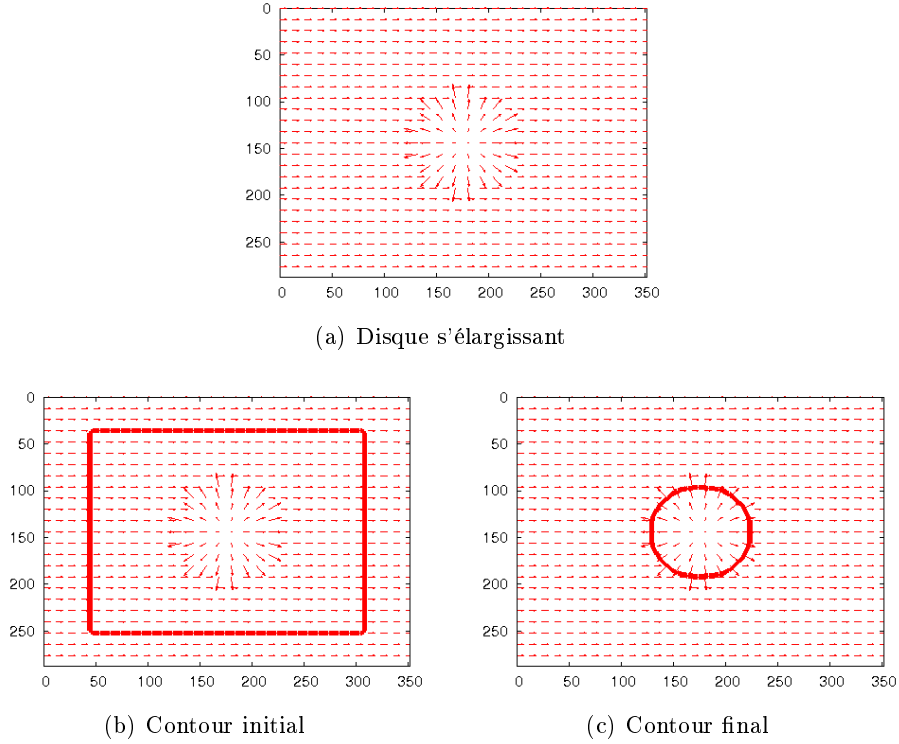


FIG. 10.4 – Disque s'élargissant : Segmentation

chaque pixel un vecteur mouvement $(u, v)^T$. Pour ce calcul, une hypothèse est faite sur la constance de la luminosité, c'est-à-dire que les objets de l'image ont une intensité constante au cours du temps. Cela peut s'écrire sous la forme de l'équation de contraintes du flot optique :

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} = 0 \quad (10.7)$$

Résoudre cette équation est un problème mal posé, car nous ne disposons que d'une équation scalaire pour chaque point de l'image où nous devons calculer un vecteur à 2 composantes. Cela implique le rajout d'une seconde contrainte qui assure d'avoir une évolution régulière du flot optique dans l'espace. Nous cherchons donc le flot optique (u, v) qui minimise la fonctionnelle suivante :

$$E(u, v) = \int_{\Omega} \left(\left(\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} \right)^2 + \alpha \Psi(|\nabla u|^2 + |\nabla v|^2) \right) dx dy \quad (10.8)$$

où Ψ est une fonction croissante différentiable, $\nabla = (\partial_x, \partial_y)^T$ l'opérateur gradient et α est un paramètre de régularisation. Cela satisfait nécessairement les équations d'Euler-

Lagrange :

$$0 = \nabla \cdot (\Psi'(|\nabla u|^2 + |\nabla v|^2) \nabla u) - \frac{1}{\alpha} \frac{\partial I}{\partial x} \left(\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} \right) \quad (10.9)$$

$$0 = \nabla \cdot (\Psi'(|\nabla u|^2 + |\nabla v|^2) \nabla v) - \frac{1}{\alpha} \frac{\partial I}{\partial y} \left(\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} \right) \quad (10.10)$$

où Ψ' est la dérivée de Ψ et $\nabla \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \partial_x a + \partial_y b$.

Une solution à ce problème peut être obtenue en plongeant les équations dans le schéma dynamique suivant :

$$u_k = \nabla \cdot (\Psi'(|\nabla u|^2 + |\nabla v|^2) \nabla u) - \frac{1}{\alpha} \frac{\partial I}{\partial x} \left(\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} \right) \quad (10.11)$$

$$v_k = \nabla \cdot (\Psi'(|\nabla u|^2 + |\nabla v|^2) \nabla v) - \frac{1}{\alpha} \frac{\partial I}{\partial y} \left(\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} \right) \quad (10.12)$$

où k est le paramètre d'évolution. Pour $k \rightarrow \infty$, la solution (u, v) représente un minimum de $E(u, v)$.

Le choix de la fonction Ψ influence le processus de régularisation et donc les résultats de l'estimation de mouvement. Nous choisissons la fonction qui a été considérée par Schnörr [Sch94] et Weickert [Wei98] :

$$\Psi(z^2) = \lambda^2 \sqrt{1 + z^2/\lambda^2} \quad (10.13)$$

Le paramètre λ est une constante positive qui sert de paramètre de contraste, voir [Wei98].

Le flot optique peut être calculé de façon itérative en utilisant les équations (10.11) et (10.12). Cela nous donne pour u et v à l'itération $k + 1$:

$$\begin{aligned} u_{k+1} &= u_k + \Delta k \left[\nabla \cdot (\Psi'(z) \nabla u_k) - \frac{1}{\alpha} \frac{\partial I}{\partial x} \left(\frac{\partial I}{\partial x} u_k + \frac{\partial I}{\partial y} v_k + \frac{\partial I}{\partial t} \right) \right] \\ v_{k+1} &= v_k + \Delta k \left[\nabla \cdot (\Psi'(z) \nabla v_k) - \frac{1}{\alpha} \frac{\partial I}{\partial y} \left(\frac{\partial I}{\partial x} u_k + \frac{\partial I}{\partial y} v_k + \frac{\partial I}{\partial t} \right) \right] \end{aligned}$$

où $z = |\nabla u_k|^2 + |\nabla v_k|^2$ et Δk est la taille du pas. Nous partons simplement d'un champ de vecteurs nuls et nous ajustons de façon itérative les vecteurs mouvement pour chaque pixel de l'image jusqu'à atteindre un critère de convergence. Le critère de convergence est souvent basé sur la différence de l'énergie à minimiser entre deux itérations.

Un problème pratique qui arrive quand on estime le flot optique est le fait que les grands déplacements ne sont pas bien estimés : l'algorithme s'arrête alors dans des minima locaux. En fait cela arrive quand les déplacements sont plus grands que la taille du masque utilisé pour l'approximation du gradient. Pour résoudre cela, une approche *multirésolution* a été implémentée par Mémin et Pérez [MP98a]. Cela signifie que le calcul du flot optique commence sur une image à une faible (sous-échantillonnée) résolution I^j et continue en augmentant (doublant) la résolution pas à pas jusqu'à obtenir la taille originale de l'image I^0 (j va de J à 0 avec J représentant la résolution la plus grossière et 0 la résolution la plus fine). A chaque résolution les vecteurs mouvement de l'étape précédente $\mathbf{u}^{j-1}$ sont

projetés sur la nouvelle résolution ($\mathbf{u}^j$) et seulement les différences $d\mathbf{u}^j$ sont calculées, c'est-à-dire que l'estimation existante est affinée. La projection (notée T ci-après) peut être une duplication ou bien une interpolation bilinéaire.

Par conséquent, la fonctionnelle (10.8) devient la suivante :

$$E(u, v) = \int_{\Omega} \left(\left(\nabla \tilde{I}^j \cdot d\mathbf{u}^j + \tilde{I}_t^j \right)^2 + \alpha \Psi \left(|\nabla(Tu^{j+1} + du^j)|^2 + |\nabla(Tv^{j+1} + dv^j)|^2 \right) \right) dx dy \quad (10.14)$$

où $\tilde{I}^j = I^j(\mathbf{x} - T\mathbf{u}^{j+1}, t)$ et $\tilde{I}_t^j = I^j(\mathbf{x}, t+1) - I^j(\mathbf{x} - T\mathbf{u}^{j+1}, t)$.

L'approche *multirésolution* peut aussi permettre d'éviter d'avoir des trous dans le champ de mouvement de zones homogènes, c'est-à-dire quand $|\nabla I|$ est faible. Cela signifie que les champs de mouvement d'objets en déplacement contenant des zones homogènes sont mieux remplis mais cela rend leurs frontières plus floues.

Le mouvement d'objets est difficile à estimer si les séquences vidéo ont été réalisées par une caméra qui se déplace. De meilleurs résultats sont obtenus en compensant le mouvement de la caméra, c'est-à-dire en estimant le mouvement global de la séquence. Pour plus de détails sur la compensation de la caméra, voir l'annexe C.

10.3.2 Résultats expérimentaux

Appliquons maintenant la méthode de segmentation par contours actifs sur les champs de vecteurs du flot optique. Les résultats suivants sont tous obtenus en utilisant des descripteurs basés sur l'entropie conditionnelle. Nous utilisons la compétition de régions entre la région comprise à l'intérieur de la courbe Ω_{in} , appelée objet, et la région comprise à l'extérieur Ω_{out} , appelée fond, de façon à ce que l'entropie des deux régions soit minimisée.

Le flot optique aussi bien que le contour actif, est calculé à deux résolutions différentes (que nous noterons *petite résolution* et *grande résolution* dans les résultats). Les caractéristiques utilisées sont la longueur du contour $|\mathbf{u}|$ et les coordonnées du vecteur (u, v) . L'utilisation de la longueur du vecteur et de sa direction n'ont pas permis d'obtenir des résultats satisfaisants, de plus l'estimation de l'angle de la direction conduit à de nombreuses erreurs.

La figure 10.5 montre le contour actif final sur la séquence **Tennis player** en utilisant la longueur du vecteur mouvement comme caractéristique pour la segmentation ($\mathbf{f}(\mathbf{x}) = |\mathbf{u}|$). Le paramètre de lissage α vaut 4000, et le paramètre de contraste $\lambda = 0.001$. Nous pouvons remarquer qu'un pied du joueur de tennis n'a pas été segmenté correctement. L'explication à cela est que le pied ne bouge pas entre les deux images successives utilisées. De façon claire, comme seulement des caractéristiques provenant du flot optique ont été utilisées, les objets qui ne bougent pas ne sont pas segmentés.

Les irrégularités de la courbe sont dues principalement aux inconsistances de l'estimation du flot optique. Il serait probablement utile d'appliquer une régularisation vectorielle afin de prendre en compte que les données sont des vecteurs et non des données vectorielles.



(a) Courbe initiale (petite résolution) (b) Itération 680 (petite résolution)



(c) Contour final (grande résolution)

FIG. 10.5 – Segmentation en utilisant la longueur du vecteur.

De plus, le petit espace entre la courbe et le côté droit de l'objet est causé par une erreur de détection de mouvement de la méthode du flot optique à des endroits qui ont été cachés (problème d'occultation).

La figure 10.6 montre le contour final en utilisant l'entropie des coordonnées des vecteurs mouvement (u, v) à la place de leur longueur. La figure 10.7 présente l'évolution des histogrammes respectifs, du fond et de l'objet. Les valeurs sont quantifiées en utilisant une grille de 20 par 20, mais seulement les parties significatives sont présentées afin de plus de clarté. Ces résultats sont presque similaires à ceux utilisant la longueur des vecteurs (on gagne un peu de précision sur le côté du joueur).

Nous testons ensuite la méthode sur la séquence **Taxi**. La figure 10.8 illustre les champs de vecteurs. La figure 10.9 montre les résultats de segmentation sur la séquence **Taxi**. Ici, considérer uniquement la longueur des vecteurs ne permet pas d'obtenir des résultats satisfaisants, en particulier sur la voiture à gauche tandis que le fait de considérer les coordonnées (u, v) donne de bien meilleurs résultats.



(a) Courbe initiale (petite résolution) (b) Courbe intermédiaire (petite résolution)



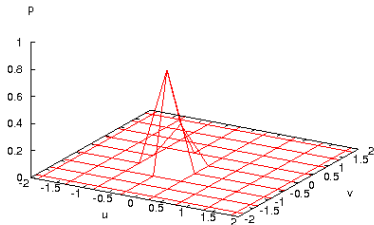
(c) Contour final (grande résolution)

FIG. 10.6 – Segmentation en utilisant les coordonnées du vecteur (u, v) .

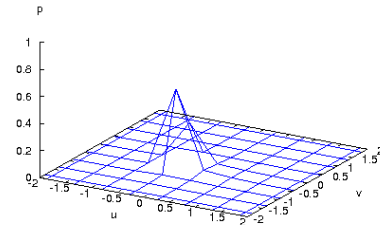
10.4 CONCLUSION

La segmentation d'objets en mouvement peut se faire en calculant un champ dense de vecteurs mouvement sur toute l'image, c'est-à-dire en calculant un vecteur mouvement par pixel. Nous avons utilisé l'entropie jointe comme critère pour la segmentation de données vectorielles, ici les vecteurs mouvement. En considérant l'entropie jointe des coordonnées des vecteurs mouvement, nous cherchons à segmenter les régions où les vecteurs mouvement sont les plus homogènes possible. De cette façon nous avons séparé les objets ayant le plus fort mouvement du reste de l'image.

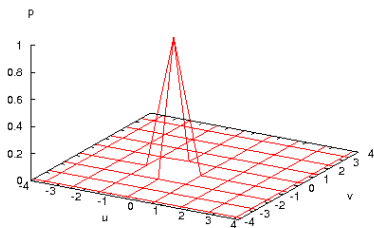
Plusieurs améliorations pourraient être apportées à la méthode. En effet, une bonne part de la qualité de la segmentation repose sur la précision de l'estimation du flot optique. Il faudrait donc obtenir une estimation du flot optique la plus robuste possible. Il serait intéressant aussi d'obtenir une méthode qui puisse gérer les occultations.



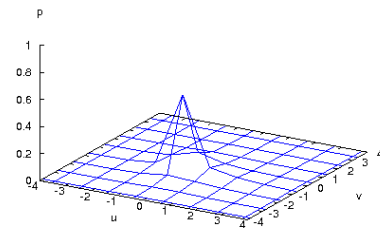
(a) Histogramme initial du fond



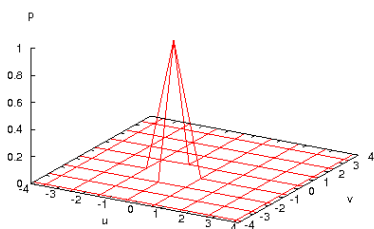
(b) Histogramme initial de l'objet



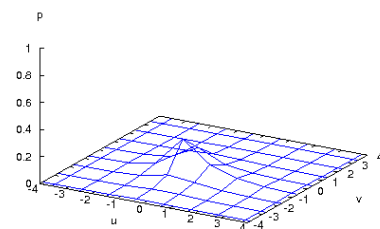
(c) Histogramme du fond, itération 120



(d) Histogramme de l'objet, itération 120

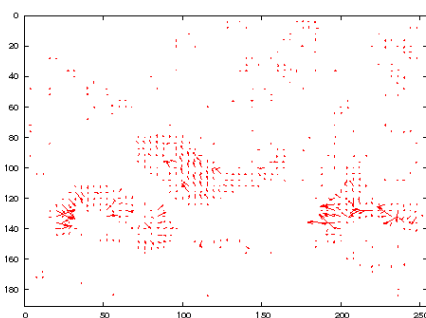


(e) Histogramme final du fond



(f) Histogramme final de l'objet

FIG. 10.7 – Evolution des histogrammes en utilisant les coordonnées du vecteur (u, v) .

FIG. 10.8 – Flot optique sur la séquence **Taxi**

(a) Contour initial (petite résolution)



(b) Contour final en utilisant la longueur du vecteur mouvement

(c) Contour final en utilisant (u, v) FIG. 10.9 – Segmentation de la séquence **Taxi** en utilisant la longueur du vecteur (à gauche) et les coordonnées du vecteur (u, v) (à droite).

Segmentation spatio-temporelle en minimisant l'entropie jointe du résiduel et de l'intensité

Dans ce chapitre, nous présentons une méthode de segmentation d'objets en mouvement, dans des séquences vidéo, qui utilise l'entropie comme critère de segmentation. Nous effectuons une segmentation spatio-temporelle des séquences vidéo en tenant compte non seulement du mouvement des objets mais aussi de leur couleur.

Tout d'abord nous présentons l'estimation de mouvement puis proposons un critère de segmentation basé sur l'erreur de prédiction. Nous expliquons ensuite la nécessité de rajouter de l'information spatiale et nous présentons le critère spatio-temporel de segmentation, basé sur l'entropie jointe. Nous intégrons ce critère dans un schéma de contours actifs et nous appliquons cette méthode à la segmentation de séquences vidéo. Nous comparons notre méthode non-paramétrique à un critère faisant des hypothèses sur les distributions des données pour valider l'intérêt d'une telle approche non-paramétrique.

11.1 POSITION DU PROBLÈME

Le mouvement d'un domaine Ω peut être calculé en choisissant un modèle de mouvement et en trouvant les paramètres de mouvement qui minimisent une fonction de l'erreur de prédiction sur le domaine. Au niveau pixel, en faisant l'hypothèse de constance de l'intensité lumineuse, l'erreur de prédiction est le résiduel suivant

$$\mathbf{e}_n(\mathbf{v}(\mathbf{x}), \mathbf{x}) = \mathbf{I}_n(\mathbf{x}) - \mathbf{I}_{n+1}(\mathbf{x} + \mathbf{v}(\mathbf{x})) \quad (11.1)$$

où $\mathbf{x}$ est un pixel du domaine Ω , $\mathbf{I}_n$ est la $n^{\text{ième}}$ image en niveaux de gris ou en couleurs de la séquence, $\mathbf{v}(\mathbf{x})$ est le mouvement apparent entre $\mathbf{I}_n$ et $\mathbf{I}_{n+1}$ au pixel $\mathbf{x}$ (appelé flot optique). Idéalement, $\mathbf{e}_n(\mathbf{x}, \mathbf{v}(\mathbf{x}))$ est égal à zéro. En niveaux de gris, il y a une seule équation pour deux inconnues (les coordonnées de $\mathbf{v}(\mathbf{x})$). Plus généralement, dans les cas niveaux de gris comme couleurs, il y a plus d'inconnues que d'équations. Par conséquent, le problème de l'estimation de mouvement doit être contraint. Une solution est de supposer que le mouvement suit un modèle défini sur Ω [WK93].

Soit $\mathbf{v}$ le vecteur des paramètres de mouvement relatifs au modèle [OB95]. Pour plus de simplicité, considérons un modèle de translation. Alors l'équation (11.1) devient :

$$\mathbf{e}_n(\mathbf{v}, \mathbf{x}) = \mathbf{I}_n(\mathbf{x}) - \mathbf{I}_{n+1}(\mathbf{x} + \mathbf{v}), \quad \mathbf{x} \in \Omega. \quad (11.2)$$

Le mouvement estimé $\hat{\mathbf{v}}$ peut être calculé comme

$$\hat{\mathbf{v}} = \arg \min_{\mathbf{v}} \sum_{\mathbf{x} \in \Omega} \varphi(\mathbf{e}_n(\mathbf{v}, \mathbf{x})) \quad (11.3)$$

ou, dans un cadre continu,

$$\hat{\mathbf{v}} = \arg \min_{\mathbf{v}} \int_{\Omega} \varphi(\mathbf{e}_n(\mathbf{v}, \mathbf{x})) \, d\mathbf{x} \quad (11.4)$$

où φ peut être, par exemple, la fonction carrée, la fonction valeur absolue, ou une fonction typique de l'estimation robuste du mouvement [Hub81, BA96].

La région issue de la segmentation basée mouvement de l'image $\mathbf{I}_n$ peut être formulée comme étant la plus grande région Ω à l'intérieur de laquelle le mouvement est cohérent avec le modèle, c'est-à-dire :

$$\begin{cases} \hat{\Omega} = \arg \min_{\Omega} \int_{\Omega} \varphi(\mathbf{e}_n(\mathbf{v}(\Gamma), \mathbf{x})) \, d\mathbf{x} \\ \mathbf{v}(\Gamma) = \arg \min_{\mathbf{v}} \int_{\Omega} \varphi(\mathbf{e}_n(\mathbf{v}, \mathbf{x})) \, d\mathbf{x} \end{cases} \quad (11.5)$$

où Γ est la frontière $\partial\Omega$ de Ω . Remarquons qu'écrire $\mathbf{v}(\Gamma)$ ou $\mathbf{v}(\Omega)$ est seulement une question de notation car Ω est complètement déterminé par Γ et réciproquement. Notons E_t , où t signifie *temporel*, l'énergie de domaine suivante :

$$E_t(\Gamma) = \int_{\Omega} \varphi(\mathbf{e}_n(\mathbf{v}(\Gamma), \mathbf{x})) \, d\mathbf{x}. \quad (11.6)$$

Le choix de φ revient à faire une hypothèse sur la distribution du résiduel $\mathbf{e}_n$ dans Ω . Par exemple, si φ est égal à la fonction carrée, l'estimation du mouvement est réalisée en faisant l'hypothèse que la distribution est gaussienne ; si φ est égal à la fonction valeur absolue, la distribution est supposée être laplacienne. Cependant, ces hypothèses ne sont en général pas vérifiées. En particulier, la présence d'*outliers* dans le résiduel, dus, par exemple, à des occultations, à une mauvaise correspondance entre le modèle de mouvement choisi et le mouvement effectif, à des variations de luminance ... peut conduire à une distribution complexe et multimodale. Par conséquent, l'estimateur de mouvement dans (11.5) peut être biaisé et amener à une perte de précision de la segmentation. De plus, un algorithme de segmentation basé sur le mouvement seulement peut échouer dans les zones homogènes ou insuffisamment texturées.

11.2 ENERGIE PROPOSÉE

Dans cette section, nous présentons l'énergie que nous utilisons dans notre algorithme. Tout d'abord nous introduisons une énergie temporelle non-paramétrique, basée sur l'entropie. Nous expliquons les modèles de mouvement utilisés pour l'estimation de mouvement et nous mettons en avant la nécessité de rajouter de l'information spatiale à cette énergie temporelle.

11.2.1 Energie temporelle

11.2.1.1 Energie non-paramétrique basée sur l'entropie

Pour prendre en compte la vraie distribution du résiduel $\mathbf{e}_n$, nous proposons de faire dépendre l'énergie de domaine E_t de l'estimation p_t de la distribution du résiduel plutôt que du résiduel lui-même. Des approches non-paramétriques ont déjà été proposées notamment dans [MP04, Com03, CM99].

L'énergie proposée est donc :

$$E_t(\Gamma) = \int_{\mathbb{R}} \psi(p_t(r)) \, dr \quad (11.7)$$

où p_t est une estimation de la distribution de $\mathbf{e}_n(\mathbf{v}(\Gamma))$ dans Ω et $\mathbf{v}(\Gamma)$ est égal à :

$$\mathbf{v}(\Gamma) = \arg \min_w \int_{\mathbb{R}} \psi(\bar{p}_t(r)) \, dr \quad (11.8)$$

avec $\bar{p}_t$ une estimation de la distribution de $\mathbf{e}_n(w)$ dans Ω . Remarquons que le domaine de définition de la distribution du résiduel est $\mathbb{R}$ en niveaux de gris (comme illustré dans (11.7) pour simplifier) ou $\mathbb{R}^3$ en couleur.

Nous proposons de choisir ψ de façon à ce que (11.7) soit la version continue de l'entropie de Shannon du résiduel :

$$\psi(p_t) = -p_t \log(p_t) . \quad (11.9)$$

L'entropie est une mesure de dispersion. Si la segmentation basée mouvement est optimale, le résiduel devrait être proche de zéro avec une dispersion minimale. De plus, l'entropie coïncide localement avec la vraisemblance à son optimum. Notons que le maximum de vraisemblance est optimal quand la distribution des données est paramétrique. Un critère de minimum d'entropie devrait donc avoir les mêmes performances qu'un critère de maximum de vraisemblance dans des cas semblables en étant capable de s'adapter à une distribution non-paramétrique. En particulier, l'entropie apparaît moins sensible aux *outliers* en pratique. L'estimation explicite de p_t n'est pas nécessaire pour calculer l'entropie [GLMI05]. Mais comme cela sera montré dans le calcul de l'équation d'évolution du contour actif, nous aurons besoin de formuler p_t explicitement. Nous utilisons une méthode de Parzen pour estimer cette distribution :

$$f_t(r) = \frac{1}{|\Omega|} \int_{\Omega} K_h(\mathbf{e}_n(\mathbf{v}(\Gamma), \mathbf{x}) - r) \, d\mathbf{x} \quad (11.10)$$

où $|\Omega|$ est le cardinal de Ω et K_h est un noyau gaussien. Il est habituel d'adapter le paramètre de lissage h aux données [Sil86, Sco92].

L'énergie proposée (11.7) est une intégrale sur $\mathbb{R}$ tandis que l'énergie classique (11.6) est une intégrale sur Ω . Pour unifier les deux approches dans un schéma commun, l'expression de l'entropie est remplacée par l'approximation suivante :

$$E_t(\Gamma) = \int_{\Omega} \psi(p_t(\mathbf{e}_n(\mathbf{v}(\Gamma), \mathbf{x}))) \, d\mathbf{x} \quad (11.11)$$

où

$$\psi(p_t) = -\frac{1}{|\Omega|} \log(p_t) . \quad (11.12)$$

Par conséquent, nous pouvons dire que l'approche paramétrique (11.6) est étendue aux distributions non-paramétriques en changeant $\varphi(\mathbf{e}_n)$ par $\psi(p_t)$, ce qui conduit à l'énergie suivante :

$$\begin{cases} E_t(\Gamma) = -\frac{1}{|\Omega|} \int_{\Omega} \log p_t(\mathbf{e}_n(\mathbf{v}(\Gamma), \mathbf{x})) \, d\mathbf{x} \\ \mathbf{v}(\Gamma) = \arg \min_w -\frac{1}{|\Omega|} \int_{\Omega} \log p_t(\mathbf{e}_n(w, \mathbf{x})) \, d\mathbf{x} \end{cases} . \quad (11.13)$$

Le domaine $\hat{\Omega}$ minimisant (11.13) représente une segmentation basée-mouvement de la frame $\mathbf{I}_n$.

11.2.1.2 Estimation du mouvement

Comme mentionné plus tôt, le mouvement $\mathbf{v}$ doit suivre un modèle dans la région Ω . Par exemple, il peut être défini par un ensemble de paramètres $\mathbf{q}$ [OB95]. Puis, estimer $\mathbf{v}$ dans (11.13) revient juste à estimer $\mathbf{q}$. C'est une tâche d'autant plus facile si la relation entre $\mathbf{v}$ et $\mathbf{q}$ est linéaire.

$$\mathbf{v}(\Gamma) = M \, \mathbf{q}(\Gamma) \quad (11.14)$$

où M est une matrice de dimension $2 \times l$ si $\mathbf{q}$ est un vecteur de dimension l . Le modèle de mouvement doit être complexe pour pouvoir prendre en compte le mouvement d'objets articulés par exemple. De ce fait, résoudre de tels problèmes est un problème inverse mal posé. Plutôt que de faire intervenir de tels modèles de mouvement complexes, nous proposons de garder le modèle simple et d'estimer les paramètres localement. L'image $\mathbf{I}_n$ est divisée en k blocs B_i de taille identique, où k dépend de la taille de l'image. Soit Ω_i l'intersection de Ω avec B_i et soit Γ_i la frontière de Ω_i . L'énergie temporelle (11.13) est remplacée par :

$$E_t^{\text{local}}(\Gamma) = -\frac{1}{|\Omega|} \int_{\Omega} \log p_t(e_n(\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{x})) \, d\mathbf{x} \quad (11.15)$$

où $\mathbf{v}_i$ est une notation pour $\mathbf{v}(\Gamma_i)$, le mouvement de Ω_i . Dans ce contexte, le modèle de mouvement peut être une simple translation. Alors l'équation (11.1) est remplacée par :

$$e_n(\mathbf{v}_i, \mathbf{x}) = \mathbf{I}_n(\mathbf{x}) - \mathbf{I}_{n+1}(\mathbf{x} + \mathbf{v}_i), \, \mathbf{x} \in \Omega_i . \quad (11.16)$$

Cette approche locale va être utilisée quand l'objet d'intérêt est articulé. Dans les autres expérimentations, une translation globale sera utilisée. Cela correspond à décomposer l'image $\mathbf{I}_n$ en un seul bloc B_1 couvrant la totalité de l'image. Pour un souci de clarté, nous noterons E_t et $e_n(\mathbf{v}(\Gamma), \mathbf{x})$ à la place de E_t^{local} et $e_n(\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{x})$, respectivement.

Finalement, pour minimiser l'influence des occultations, nous utilisons une expression différente pour (11.2) :

$$e_n(\mathbf{v}, \mathbf{x}) = \min_{\text{abs}} \{ \mathbf{I}_n(\mathbf{x}) - \mathbf{I}_{n+1}(\mathbf{x} + \mathbf{v}), \mathbf{I}_n(\mathbf{x}) - \mathbf{I}_{n-1}(\mathbf{x} - \mathbf{v}) \} \quad (11.17)$$

où $\min_{\text{abs}}$ est égale à :

$$\min_{\text{abs}}\{a, b\} = \begin{cases} a & \text{si } \min\{|a|, |b|\} = |a| \\ b & \text{si } \min\{|a|, |b|\} = |b| \end{cases} . \quad (11.18)$$

La fonction $\min_{\text{abs}}$ n'est pas différentielle mais nous n'avons pas besoin de la dériver.

11.2.2 Nécessité de considérer l'information spatiale

L'énergie (11.13) est bien adaptée pour la segmentation d'objets sur un fond texturé. Mais il peut arriver que la segmentation englobe des zones homogènes ou quasi-homogènes. En effet, de telles zones ont un faible résiduel si elles sont compensées avec le mouvement estimé de l'objet, du moins tant que l'objet compensé en mouvement reste dans une zone homogène. Donc, l'énergie (11.13) devrait augmenter de façon négligeable en rajoutant de telles zones. De plus, comme les notions d'objet et de fond sont arbitraires et peuvent être échangées, nous pouvons remarquer que ce phénomène peut aussi se produire si l'objet contient des zones homogènes près de sa frontière.

La solution proposée est de combiner de l'information spatiale et de l'information temporelle [BRDW03]. Pour se conformer au schéma proposé, l'entropie va devenir une entropie jointe du résiduel et de l'intensité ou la couleur de l'objet. Intuitivement, l'entropie de la couleur de l'objet va augmenter si le domaine de l'objet inclut du fond car de nouvelles couleurs vont être ajoutées à l'objet. Si le fond a la même couleur que l'objet près de sa frontière, alors il n'y a pas d'information objective pour trouver la frontière de l'objet. De plus, cela augmente la dispersion de la distribution de la couleur. Par conséquent, cela fait augmenter l'entropie.

La figure 11.1 illustre ce phénomène. Cette séquence vidéo est filmée depuis un véhicule en déplacement. Toute la scène a donc un mouvement. Seulement par effet de perspective les objets du premier plan comme l'arbre ont un mouvement beaucoup plus prononcé que le fond de la scène. Nous cherchons à segmenter l'arbre qui a un mouvement apparent de la droite vers la gauche. Sur cette image, nous montrons une segmentation correcte de l'arbre (colonne de gauche) et les *pdf* respectivement du résiduel et de l'intensité lumineuse (au centre et en bas). Sur la colonne de droite sont représentées les mêmes illustrations en incluant une partie du ciel (homogène) dans la segmentation de l'objet. Nous pouvons remarquer qu'inclure une partie du ciel ne fait que rajouter des zéros dans la *pdf* du résiduel et donc ne fait pas augmenter de façon significative l'entropie du résiduel. Au contraire en considérant l'intensité lumineuse de l'objet, rajouter une partie du fond fait apparaître une nouvelle modalité dans la *pdf* de l'objet et fait donc augmenter l'entropie de l'intensité de l'objet. Cette illustration montre qu'il peut être intéressant de considérer l'information spatiale dans le critère de segmentation des objets en mouvement, pour éviter de sur-segmenter en présence de zones homogènes.

Nous pouvons remarquer que cette approche définit un schéma général pour de la segmentation multimodale : l'entropie conjointe permet de combiner un nombre arbitraire



(a) Segmentation de l'arbre



(b) Segmentation de l'arbre en englobant du ciel

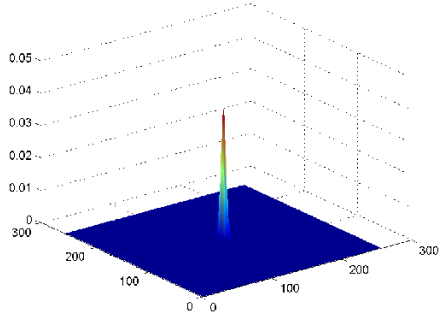
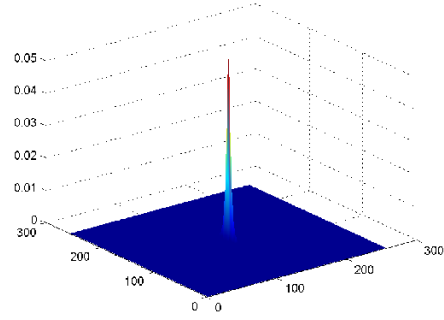
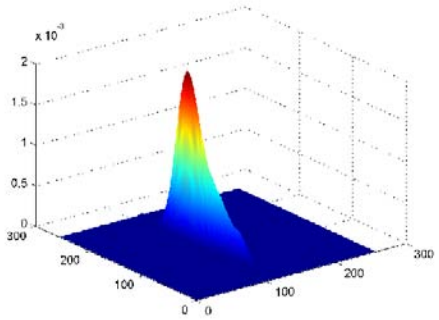
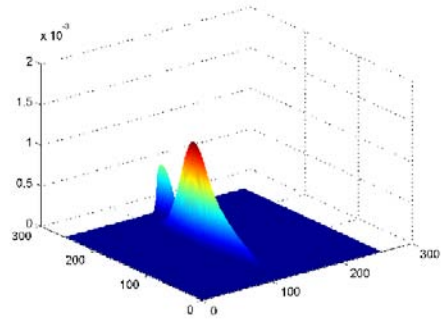
(c) Pdf du résiduel $\mathbf{e}_n$. Entropie : 4.65(d) Pdf du résiduel $\mathbf{e}_n$. Entropie : 4.50(e) Pdf de l'intensité $\mathbf{I}_n$. Entropie : 7.21(f) Pdf de l'intensité $\mathbf{I}_n$. Entropie : 7.54

FIG. 11.1 – Comparaison des fonctions de densité de probabilité (*pdf*) du résiduel et de l'intensité de l'image à l'intérieur du contour en segmentant correctement l'arbre (à gauche) et en incluant une partie du fond homogène (à droite). La *pdf* du résiduel a été ramenée de $[-255; 255]$ à $[0; 255]$ pour une meilleure lisibilité et la *pdf* de l'intensité est représentée dans l'espace couleur YUV en considérant les canaux Y et U.

de modalités. En pratique, ce nombre est limité par le nombre d'échantillons disponibles, c'est-à-dire le nombre de pixels de l'image. Si les échantillons sont trop peu nombreux alors l'entropie sera mal estimée.

11.2.3 Combiner les énergies spatiale et temporelle

Comme présenté précédemment l'information spatiale et temporelle peuvent être combinées de façon cohérente en considérant l'entropie conjointe du résiduel et de l'intensité ou la couleur de l'objet. L'énergie spatio-temporelle est alors la suivante :

$$E(\Gamma) = -\frac{1}{|\Omega|} \int_{\Omega} \log p(\mathbf{e}_n(\mathbf{v}(\Gamma), \mathbf{x}), \mathbf{I}_n(\mathbf{x})) \, d\mathbf{x} \quad (11.19)$$

où p est maintenant la distribution conjointe du résiduel et de la couleur de l'image à l'intérieur du domaine de l'objet Ω .

Supposons que le résiduel et la couleur sont indépendants. Considérons le modèle de séquence suivant :

$$\mathbf{I}_{n+1}(\mathbf{x}) = \mathbf{I}_n(T(\mathbf{x})) + n(\mathbf{x}) \quad (11.20)$$

où T est une transformation et n un bruit blanc gaussien. Le résiduel vaut :

$$\mathbf{e}_n(\mathbf{v}(\mathbf{x}), \mathbf{x}) = \mathbf{I}_n(\mathbf{x}) - \mathbf{I}_{n+1}(\mathbf{v}(\mathbf{x})) . \quad (11.21)$$

si le mouvement est estimé parfaitement alors $\mathbf{v}$ est égal à T^{-1} et :

$$\begin{aligned} \mathbf{e}_n(\mathbf{v}(\mathbf{x}), \mathbf{x}) &= \mathbf{I}_n(\mathbf{x}) - \mathbf{I}_{n+1}(T^{-1}(\mathbf{x})) \\ &= \mathbf{I}_n(\mathbf{x}) - \mathbf{I}_n(\mathbf{x}) + n(T^{-1}(\mathbf{x})) \\ &= n(T^{-1}(\mathbf{x})) . \end{aligned} \quad (11.22)$$

$\mathbf{e}_n(\mathbf{v}(\mathbf{x}), \mathbf{x})$ est alors indépendant de $\mathbf{I}_n$.

Si la transformation n'est pas correctement estimée, alors le résiduel est certainement corrélé avec $\mathbf{I}_n$. Il faut alors estimer des densités de probabilité jointes et ce n'est pas simple avec une méthode à noyaux comme celle de Parzen. D'autres méthodes, comme par exemple celles basées sur les k -ièmes plus proches voisins sont plus adaptées pour travailler dans des dimensions supérieures à 1 (voir par exemple les travaux en cours de [BDB07]).

Cependant la corrélation entre $\mathbf{I}_n$ et e_n est seulement partielle. En fait, le modèle (11.20) n'est pas réaliste. Il n'existe pas de telle transformation T , l'image $\mathbf{I}_{n+1}$ étant une projection sur un plan 2D d'une scène 3D. En général quelques parties de l'objet dans $\mathbf{I}_n$ deviennent invisibles dans $\mathbf{I}_{n+1}$ tandis que d'autres deviennent visibles. Alors l'image $\mathbf{I}_{n+1}$ ne peut pas être déduite entièrement de $\mathbf{I}_n$. En d'autres termes, le résiduel contient de l'information imprévisible étant donné $\mathbf{I}_n$. Au final, si la transformation représentée par le mouvement $\mathbf{v}(\Gamma)$ est bien estimée alors il est tout de même raisonnable de faire l'hypothèse d'indépendance.

En utilisant cette hypothèse, l'énergie (11.19) peut être ré-écrite comme suit :

$$\begin{aligned} E(\Gamma) &= -\frac{1}{|\Omega|} \int_{\Omega} \log p_t(\mathbf{e}_n(\mathbf{v}(\Gamma), \mathbf{x})) \, d\mathbf{x} - \frac{1}{|\Omega|} \int_{\Omega} \log p_s(\mathbf{I}_n(\mathbf{x})) \, d\mathbf{x} \\ &= E_t(\Gamma) + E_s(\Gamma) \end{aligned} \quad (11.23)$$

où s dans E_s et p_s signifie *spatial* et $\mathbf{v}(\Gamma)$ est définie dans (11.13).

Remarquons qu'en faisant cette hypothèse d'indépendance nous obtenons une somme de deux énergies, ce qui est fréquemment utilisé quand on veut minimiser simultanément plusieurs énergies. Nous obtenons ici la somme de deux entropies. Notons que généralement des termes de pondération sont introduits pour mesurer l'influence de chaque énergie, ce qui n'est ici pas le cas et nous n'avons donc pas de paramètre à régler entre les termes. Ce paramètre n'est en général pas facile à régler et est souvent déterminé expérimentalement.

11.3 VERS L'ÉQUATION D'ÉVOLUTION DU CONTOUR ACTIF

Pour minimiser l'énergie (11.23), nous devons calculer sa dérivée par rapport à Γ . Nous pouvons montrer que la dérivée de forme dE_t de E_t est donnée par :

$$\begin{aligned} dE_t(\Gamma, F) &= \frac{1}{|\Omega|} \int_{\Gamma} \left[\log p_t(\mathbf{e}_n(\mathbf{v}(\Gamma), s)) - 1 + E_t(\Gamma) \right. \\ &\quad \left. + \frac{1}{|\Omega|} \int_{\Omega} \frac{K_{\sigma}(\mathbf{e}_n(\mathbf{v}(\Gamma), s) - \mathbf{e}_n(\mathbf{v}(\Gamma), \mathbf{x}))}{p_t(\mathbf{e}_n(\mathbf{v}(\Gamma), \mathbf{x}))} \, d\mathbf{x} \right] N(s) \cdot F(s) \, ds \end{aligned} \quad (11.24)$$

où N est la normale intérieure de Γ .

Nous pouvons remarquer que la distribution p_t apparaît explicitement dans la dérivée comme mentionné, dans la partie 11.2.1.

L'expression de $dE_s(\Gamma, F)$ est similaire. La dérivée de forme de (11.23) est égale à :

$$dE(\Gamma, F) = dE_t(\Gamma, F) + dE_s(\Gamma, F) \quad (11.25)$$

La dérivée (11.24) a donc la forme suivante :

$$dE_t(\Gamma, F) = \int_{\Gamma} (\alpha_t(s) N(s)) \cdot F(s) \, ds = \langle \alpha_t N, F \rangle_{L_2} \quad (11.26)$$

où $\langle \cdot, \cdot \rangle_{L_2}$ est le produit scalaire L_2 sur Γ . Alors, $\alpha_t N$ est par définition le gradient de (11.13) à Γ . Le gradient de E_s est obtenu de la même façon et mis sous la forme $\alpha_s N$.

A partir de la dérivée de l'énergie nous pouvons en déduire l'équation d'évolution du contour actif :

$$\begin{cases} \Gamma(\tau = 0) = \Gamma_0 \\ \frac{\partial \Gamma}{\partial \tau} = -(\alpha_t + \alpha_s) N \end{cases} \quad (11.27)$$

où τ est le paramètre d'évolution. Classiquement, une descente de gradient est effectuée jusqu'à obtenir la condition d'arrêt $\alpha_t + \alpha_s = 0$.

Nous choisissons de considérer l'entropie de l'intensité lumineuse sachant que l'on se trouve à l'intérieur ou à l'extérieur du contour. Pour cela notons Ω_{in} la région comprise à l'intérieur du contour (l'objet), de frontière Γ_{in} et Ω_{out} la région comprise à l'extérieur, de frontière Γ_{out} . D est le domaine de l'image. Nous définissons la fonction C de façon à ce qu'elle caractérise l'objet : $C = 1$ à l'intérieur de l'objet et $C = 0$ à l'extérieur. L'entropie jointe du résiduel et de l'intensité lumineuse conditionnellement à C s'écrit :

$$H(\mathbf{e}_n, \mathbf{I}_n | C) = \sum_{i \in \{0,1\}} p(C = i) H(\mathbf{e}_n, \mathbf{I}_n | C = i) \quad (11.28)$$

avec $p(C = 0)$ la probabilité de l'événement $C = 0$, et $p(C = 1)$ la probabilité de $C = 1$. Or l'entropie $H(\mathbf{e}_n, \mathbf{I}_n | C = 1)$ peut être approximée et s'écrire comme l'énergie $E(\Gamma_{in})$ définie en (11.19), c'est une approximation de l'entropie conjointe entre le résiduel et l'intensité, conditionnellement à $C = 1$. De la même façon, $H(\mathbf{e}_n, \mathbf{I}_n | C = 0)$ peut être approximée par $E(\Gamma_{out})$. L'énergie 11.28 s'écrit alors :

$$\begin{aligned} H(\mathbf{e}_n, \mathbf{I}_n | C) &= p(C = 1) E(\Gamma_{in}) + p(C = 0) E(\Gamma_{out}) \\ &= \frac{|\Omega_{in}|}{|D|} E(\Gamma_{in}) + \frac{|\Omega_{out}|}{|D|} E(\Gamma_{out}) \end{aligned} \quad (11.29)$$

En multipliant l'énergie par une constante sans influencer la segmentation nous obtenons :

$$\begin{aligned} H(\mathbf{e}_n, \mathbf{I}_n | C) | D| &= |\Omega_{in}| E(\Gamma_{in}) + |\Omega_{out}| E(\Gamma_{out}) \\ &= E_{rc} \end{aligned} \quad (11.30)$$

en notant E_{rc} l'énergie en compétition de régions.

Nous nous retrouvons ainsi avec une énergie en compétition de régions, qui prend en compte la taille des régions. De plus cette énergie est équivalente au critère du maximum de vraisemblance qui est fréquemment utilisé.

La dérivée de forme de (11.30) peut être obtenue en appliquant la loi de dérivation $(u v)' = u' v + u v'$, et en utilisant la dérivée suivante :

$$d(|\Omega|)(\Gamma, F) = d\left(\int_{\Omega} d\mathbf{x}\right)(\Gamma, F) \quad (11.31)$$

$$= - \int_{\Gamma} N(s) \cdot F(s) ds . \quad (11.32)$$

Les termes relatifs à l'objet et ceux relatifs au fond peuvent être assemblés en notant que Γ_{in} et Γ_{out} sont identiques à un signe d'orientation près. En particulier, la normale intérieure N_{out} de Γ_{out} est égale à $-N$.

Nous avons donc présenté l'énergie spatio-temporelle qui va servir à segmenter des objets en mouvement dans les séquences vidéo. Nous présentons maintenant quelques résultats sur différentes séquences vidéo. Nous comparons les énergies temporelle, spatiale et spatio-temporelle ainsi que notre critère non-paramétrique à un critère utilisant des distributions paramétriques.

11.4 RÉSULTATS EXPÉRIMENTAUX

Nous présentons dans cette section quelques résultats expérimentaux pour montrer l'efficacité de notre énergie spatio-temporelle. Nous rappelons tout d'abord les conditions d'expérimentations puis nous comparons les énergies temporelle, spatiale et spatio-temporelle sur des séquences synthétiques et des séquences réelles. Nous comparons ensuite notre approche à une méthode utilisant des distributions paramétriques. Finalement nous présentons des résultats en suivi d'objet en mouvement en utilisant les deux modèles de mouvement proposés : le global et le local.

11.4.1 Conditions d'expérimentations

Nous exposons les conditions d'expérimentation en donnant quelques détails sur l'implémentation.

Pour rappel, l'énergie proposée a la forme suivante :

$$E_{rc}(\Gamma) = |\Omega_{in}| E(\Gamma_{in}) + |\Omega_{out}| E(\Gamma_{out}) \quad (11.33)$$

où (voir Eq. (11.23))

$$E(\Gamma_i) = E_t(\Gamma_i) + E_s(\Gamma_i) \quad \text{avec } i \in \{in, out\} . \quad (11.34)$$

Pour effectuer des comparaisons, l'énergie (11.33) sera aussi utilisée sous deux formes incomplètes : quand E_s est supprimé de la définition de E dans (11.34), l'énergie sera appelée énergie temporelle ; quand E_t est supprimé de la définition de E , l'énergie sera appelée énergie spatiale. Dans sa forme complète, elle est appelée énergie spatio-temporelle.

Les tests ont été effectués sur des séquences synthétiques et réelles. Les séquences synthétiques sont composées d'images de taille 300×300 pixels et les séquences réelles sont des séquences en format `cif` (352×288 pixels), toutes étant représentées dans l'espace couleur YUV. Nous ne tenons pas compte du canal couleur V et considérons donc les deux canaux Y et U . Les distributions de $\mathbf{I}_n$ et $\mathbf{e}_n$ sont des fonctions de $\mathbb{R}^2$ vers $\mathbb{R}$ construites à partir d'échantillons de $[0..255]^2$ et de $[-255, 255]^2$ respectivement. (L'intervalle $[a..b]$ représente des entiers tandis que $[a, b]$ représente des réels. Les échantillons de $\mathbf{e}_n$ ne sont pas forcément des entiers car $\mathbf{I}_{n+1}(\mathbf{x} + \mathbf{v})$ est interpolée bilinéairement). Nous supposons l'indépendance entre les canaux Y et U et de ce fait les termes d'entropie s'écrivent comme la somme des entropies sur chaque composante (de la même façon que nous obtenons l'écriture de l'énergie E comme la somme des énergies E_t et E_s).

Pour estimer la densité de probabilités des données sur la région Ω , nous utilisons une méthode à noyaux. La largeur de bande de ce noyau est calculé en utilisant la déviation standard sur la région [Sil86] :

$$h = 0.9 \min(\hat{\sigma}, \hat{q}/1.34) |\Omega|^{-1/5} \quad (11.35)$$

où $\hat{\sigma}$ est la déviation standard et $\hat{q}$ l'étendue interquartile, des données estimées sur la région Ω . Cette méthode d'estimation de largeur de bande est appelée *plug-in*.

L'estimation du mouvement a été implémentée en utilisant un modèle de translation. Nous avons utilisé une méthode sous-optimale rapide (la recherche en diamant [TRRK98]) qui permet dans notre cas un mouvement de $-12/+12$ pixels dans les deux directions avec une précision au quart de pixel. Cette procédure a été utilisée pour le mouvement global sur la région Ω et pour le mouvement local sur chaque Ω_i .

Le nombre de points d'échantillonnage de la spline représentant le contour est de 60 et le pas d'évolution est de 0.5.

Dans les résultats suivants, la segmentation consiste en une initialisation loin de la solution (généralement un cercle) suivie de l'évolution du contour, le tout sur une seule image. Un suivi d'objets sur une séquence d'images consiste en une initialisation provenant de la segmentation de l'image précédente et translatée du mouvement de l'objet, suivie par l'évolution du contour actif.

11.4.2 Comparaisons avec énergie spatiale, temporelle et spatio-temporelle

Dans cette première partie nous comparons les résultats obtenus en utilisant l'énergie spatiale seule, l'énergie temporelle seule et enfin l'énergie spatio-temporelle. Nous présentons tout d'abord des segmentations de séquences synthétiques puis de séquences réelles.

11.4.2.1 Séquences synthétiques

Les séquences synthétiques suivantes ont été créées en combinant différents motifs texturés et homogènes. Un objet se déplace en translation de -3 pixels sur un fond se déplaçant de 1 pixel. La figure 11.2 présente les résultats de segmentation sur différentes combinaisons de textures, en utilisant le critère temporel seul, le critère spatial seul et le critère spatio-temporel. Les résultats tendent à montrer qu'en présence de textures, le critère temporel est efficace, en présence de zones homogènes, le critère spatial se comporte mieux que le critère temporel et dans tous les cas le critère spatio-temporel réussit à segmenter correctement l'objet. La combinaison d'informations spatiale et temporelle semble appropriée pour segmenter des séquences qui contiennent à la fois des zones homogènes et texturées.

11.4.2.2 Séquences de test réelles

Comme dans la section précédente, nous comparons l'énergie spatio-temporelle à l'énergie spatiale seule et à l'énergie temporelle seule. La première séquence est **Flowers and garden** (figure 11.3). C'est une séquence prise en caméra mobile, et donc l'arbre du premier plan a un déplacement apparent bien plus fort que les maisons du fond de l'image. La segmentation y est difficile à cause des textures des fleurs et du ciel homogène qui ne permet pas une estimation correcte du mouvement. En utilisant l'énergie temporelle seule, quand le contour atteint le ciel, l'énergie temporelle diminue, et donc nous pouvons observer une

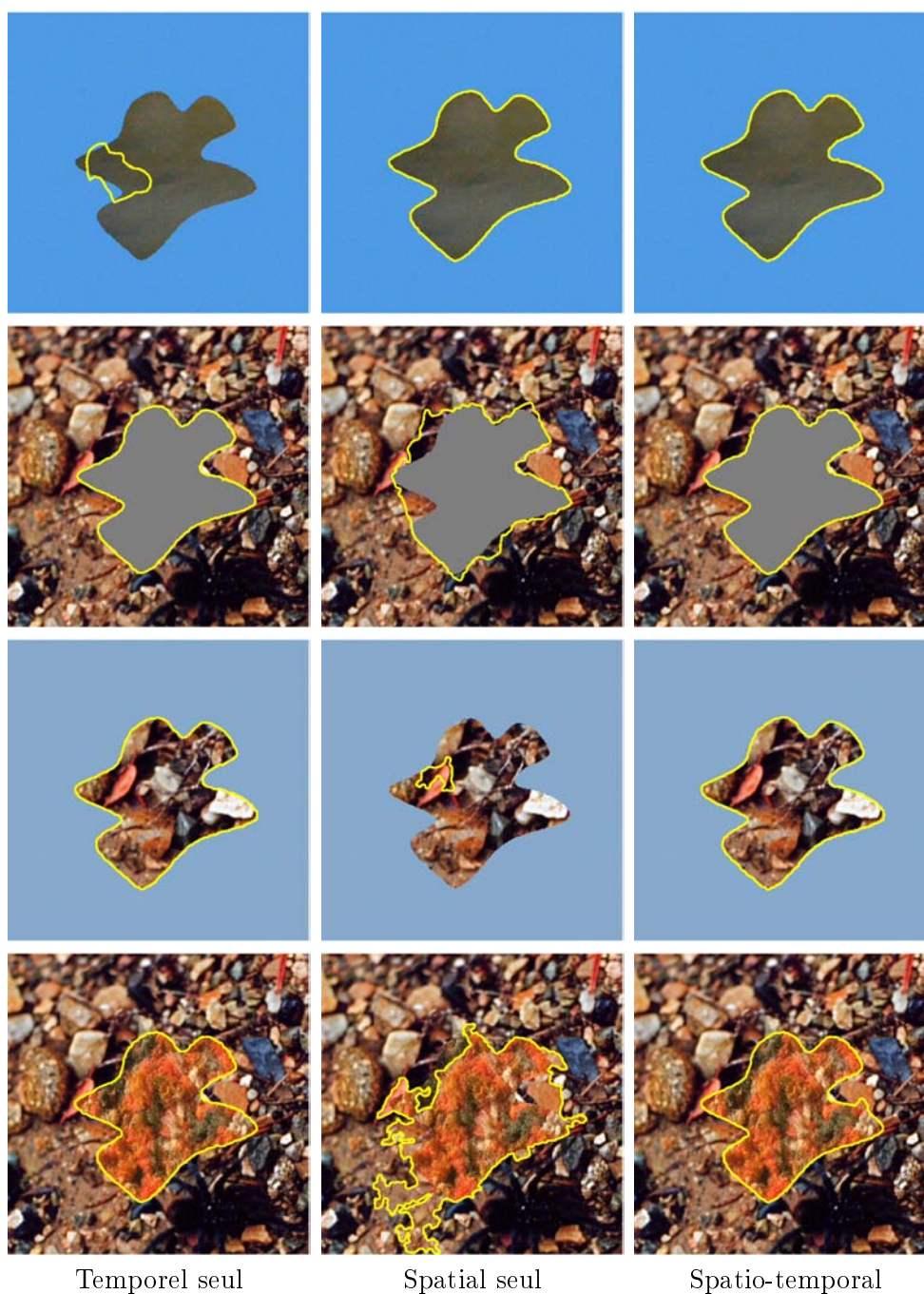


FIG. 11.2 – Comparaisons des critères temporel, spatial et spatio-temporel pour la segmentation d'objets synthétiques sur des fonds synthétiques. Première ligne, un objet homogène sur un fond homogène ; seconde ligne, un objet homogène sur un fond texturé ; troisième ligne, un objet texturé sur un fond homogène ; dernière ligne, un objet texturé sur un fond texturé.

sur-segmentation de l'arbre sur le ciel. Avec l'énergie spatiale seule, l'algorithme de segmentation inclut les maisons dans la segmentation du fait qu'elles sont de la même couleur que l'arbre. Finalement, l'énergie spatio-temporelle permet de ne pas sur-segmenter le ciel et de ne pas déborder sur les maisons.

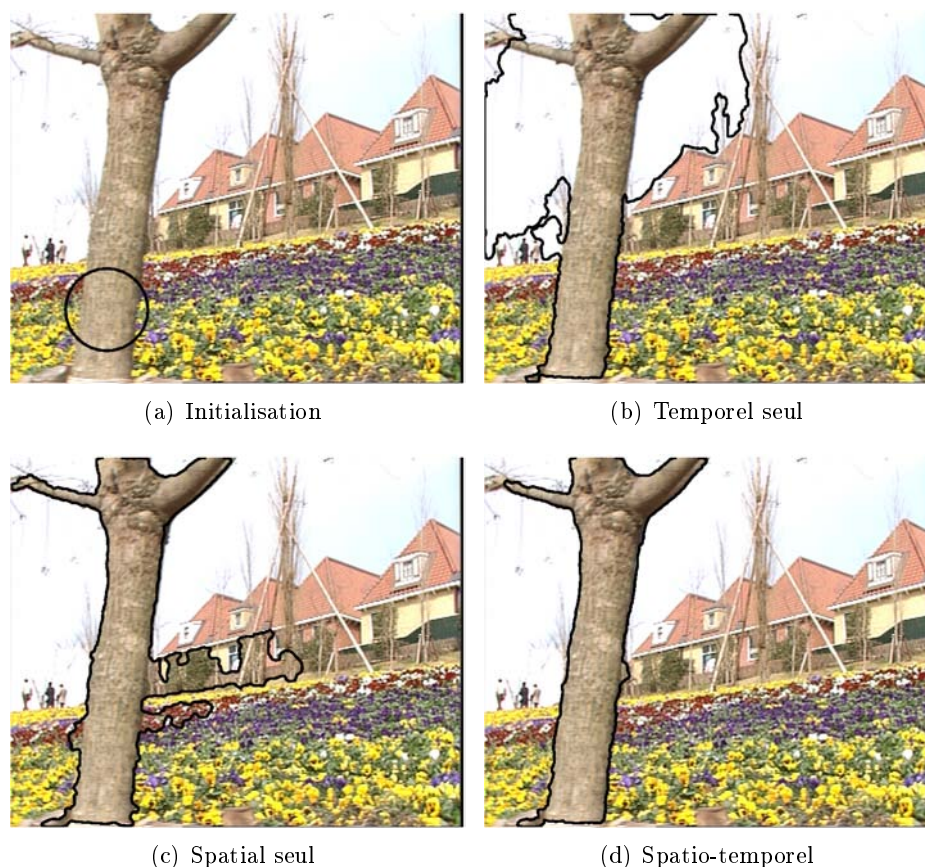


FIG. 11.3 – Segmentation de l'arbre sur la séquence 'Flowers and garden'.

La seconde séquence présentée est **Soccer** (figure 11.4). Cette séquence présente la difficulté d'avoir un objet articulé, le joueur de football, à segmenter. L'objet ne réalise pas une transformation rigide comme permise par notre modèle de mouvement. De plus, les couleurs de sa tête et de ses pieds sont proches des couleurs contenues dans le fond. L'énergie temporelle seule ne permet de segmenter que les zones qui obéissent au modèle de mouvement, tandis que l'énergie spatiale ne réussit pas à segmenter la tête à cause de sa couleur. L'énergie spatio-temporelle réussit bien à segmenter le joueur, elle rate parfois un pied du joueur quand aucune des deux énergies seules ne réussit à le segmenter. En effet, nous pouvons observer que l'énergie spatiale aide le terme temporel aux endroits qui n'obéissent pas au modèle de mouvement, c'est-à-dire aux zones articulées tandis que l'énergie temporelle aide le terme spatial quand la couleur n'est pas discriminante.

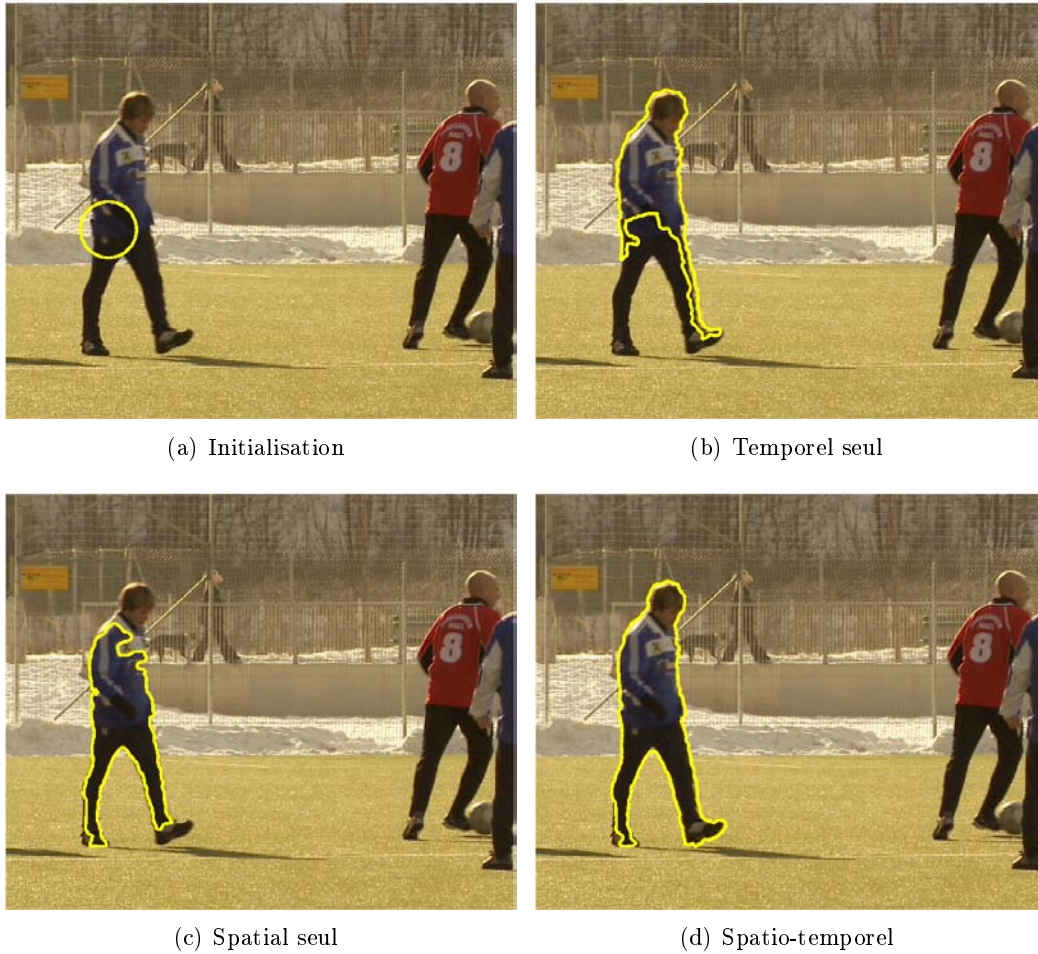


FIG. 11.4 – Segmentation d'un joueur sur la séquence 'Soccer'.

La troisième séquence est **Football** (figure 11.5). Là encore, c'est une séquence difficile de par le mouvement non-rigide qu'effectue le joueur de football tout comme les couleurs très différentes de sa tenue. De même que pour l'exemple précédent, l'énergie temporelle ne permet de segmenter que les parties qui bougent de façon rigide et l'énergie spatiale réussit à segmenter les zones les plus homogènes. La combinaison des deux parvient à réaliser une meilleure segmentation même s'il manque un pied à cause de son mouvement non rigide et de sa couleur bien distincte.

11.4.3 Approche non-paramétrique comparée à l'approche paramétrique

Nous pouvons nous demander quelles améliorations apporte le fait de considérer des estimations non-paramétriques de la distribution de l'erreur de prédiction et de celle de la couleur, par opposition aux termes classiques d'erreur correspondant à des hypothèses paramétriques. Pour une juste comparaison, les hypothèses paramétriques des distributions

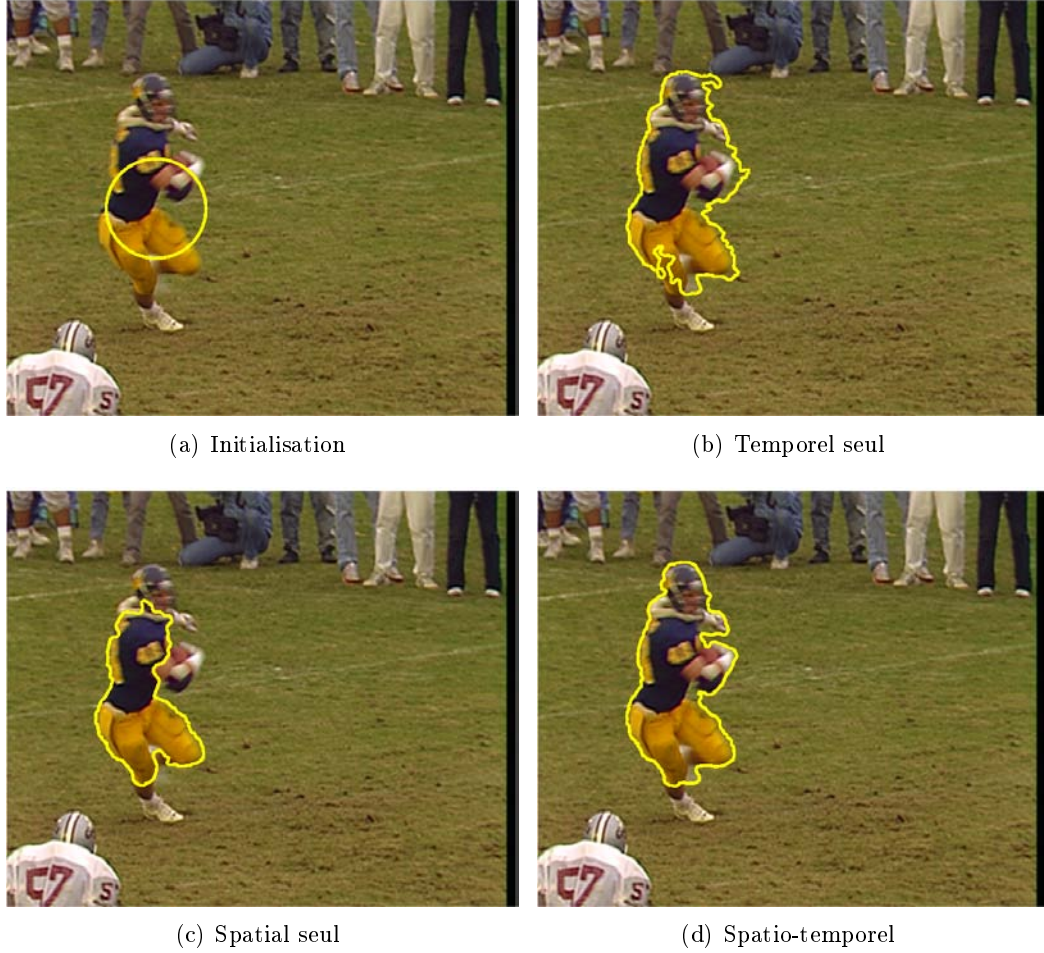


FIG. 11.5 – Segmentation d'un joueur sur la séquence 'Football'.

doivent être choisies de façon appropriée.

11.4.3.1 Comparaison temporelle

Nous comparons tout d'abord l'énergie temporelle que nous proposons à une énergie basée sur une approche paramétrique dans le sens qu'elle suppose une distribution spécifique des données.

L'erreur de prédiction est corrompue par des *outliers* dus en particulier au bruit, à des variations d'illumination, à des erreurs de correspondance avec le modèle de mouvement et à des occultations. Choisissons la somme des valeurs absolues (appelée généralement SAD) comme terme classique d'erreur [WK93], pour sa robustesse aux *outliers* et notons que ce terme provient d'une hypothèse laplacienne. En effet, si le résiduel $\mathbf{e}_n$ est un champ aléatoire spatialement décorrélé avec une distribution laplacienne de moyenne μ_e et d'échelle σ , la

probabilité d'avoir un tel champ, conditionnée par un mouvement $\mathbf{v}$, est égale à :

$$p(\mathbf{e}_n|\mathbf{v}) = \frac{1}{(2\sigma)^{|\Omega|}} \prod_{\mathbf{x} \in \Omega} \exp - \frac{|\mathbf{e}_n(\mathbf{v}, \mathbf{x}) - \mu_e|}{\sigma}. \quad (11.36)$$

L'estimation du maximum de vraisemblance de $\mathbf{v}$ est équivalente à :

$$\arg \max_{\mathbf{v}} \prod_{\mathbf{x} \in \Omega} \exp - \frac{|\mathbf{e}_n(\mathbf{v}, \mathbf{x}) - \mu_e|}{\sigma} \quad (11.37)$$

$$\iff \arg \max_{\mathbf{v}} \log \left(\prod_{\mathbf{x} \in \Omega} \exp - \frac{|\mathbf{e}_n(\mathbf{v}, \mathbf{x}) - \mu_e|}{\sigma} \right) \quad (11.38)$$

$$\iff \arg \min_{\mathbf{v}} \sum_{\mathbf{x} \in \Omega} \frac{|\mathbf{e}_n(\mathbf{v}, \mathbf{x}) - \mu_e|}{\sigma} \quad (11.39)$$

En pratique, choisir μ_e différent de zéro peut seulement être motivé par un changement global de l'illumination entre les images $\mathbf{I}_n$ et $\mathbf{I}_{n+1}$. En faisant l'hypothèse que l'illumination globale reste constante, μ_e va être mis à zéro. Alors l'estimation (11.39) est équivalente à :

$$\arg \min_{\mathbf{v}} \sum_{\mathbf{x} \in \Omega} \frac{|\mathbf{I}_n(x) - \mathbf{I}_{n+1}(\mathbf{x} + \mathbf{v})|}{\sigma} \quad (11.40)$$

qui est le critère SAD.

Pour la comparaison, nous utilisons une approximation différentiable de la SAD. Nous notons cette méthode ABS_e dans la suite.

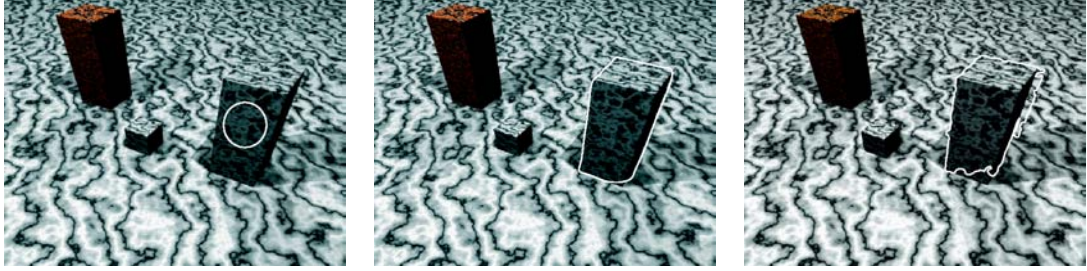
Nous évaluons l'énergie temporelle sur la séquence synthétique **“Marbled blocks”** dans laquelle plusieurs blocs évoluent sur un fond constitué de la même texture. Les blocs et le fond ayant la même texture, seul un critère temporel peut réussir à segmenter correctement les blocs. Nous choisissons de segmenter le bloc à droite dans la séquence.

La figure 11.6 montre les résultats obtenus par les deux méthodes pour la segmentation du bloc de droite en prenant un cercle comme initialisation. Le mouvement a été estimé entre deux images successives. La méthode ABS_e ne réussit pas à segmenter correctement le bloc tandis que notre méthode y parvient.

Ensuite nous présentons sur la figure 11.7 un suivi de ce bloc de l'image 50 à l'image 200. Pour avoir un fort mouvement et montrer la stabilité de notre méthode, nous avons sous-échantillonné la séquence à un rapport de 1/5. Ces images montrent que nous obtenons un meilleur suivi que la méthode ABS_e , notamment sur le côté droit où se produit un phénomène d'occultation. La méthode paramétrique est donc plus instable que la notre, surtout aux endroits où le mouvement est mal estimé.

Nous présentons enfin sur la figure 11.8 les résultats de segmentation en utilisant les deux méthodes. Sur cette séquence le bâtiment du milieu a un fort mouvement par-rapport au fond mais ce mouvement est plus complexe que la translation permise par notre modèle de mouvement. Nous pouvons toutefois constater que la segmentation est meilleure avec notre approche non-paramétrique par-rapport à la méthode classique ABS_e .

D'après les expérimentations, la segmentation semble donner de meilleurs résultats avec notre approche non-paramétrique qu'avec une approche paramétrique classique, notamment au niveau des occultations.



(a) Initialisation avec un cercle (b) Segmentation obtenue avec notre méthode (c) Segmentation obtenue avec la méthode ABS_e

FIG. 11.6 – Comparaison du terme temporel avec une méthode classique paramétrique de l'image 50 de la séquence **Marbled blocks**

11.4.3.2 Comparaison spatio-temporelle

Nous voulons maintenant comparer le critère spatio-temporel à un critère équivalent en paramétrique. Concernant l'énergie temporelle, nous reprenons ce qui est présenté dans la partie précédente. Nous étudions alors le critère spatial, nous étudions ici une équivalence paramétrique.

Il est évident qu'il n'existe pas d'hypothèse paramétrique idéale concernant le terme spatial. Cependant, en notant que le critère spatial peut être interprété comme un critère d'homogénéité de couleur par morceaux, il semble raisonnable de faire une hypothèse de distribution gaussienne. Si la couleur de l'image $\mathbf{I}_n$ est un champ aléatoire décorrélié spatialement de distribution gaussienne de moyenne μ_I et de déviation standard σ_I , la probabilité d'avoir un tel champ, conditionné à un mouvement $\mathbf{v}$, est égale à :

$$P(\mathbf{I}_n|\mathbf{v}) = \frac{1}{\sqrt{2\pi} \sigma_I} \prod_{\mathbf{x} \in \Omega} \exp - \frac{(\mathbf{I}_n(\mathbf{x}) - \mu_I)^2}{2\sigma_I^2} . \quad (11.41)$$

L'estimation du maximum de vraisemblance de $\mathbf{v}$ est alors équivalent à minimiser la somme des différences au carré (SSD) :

$$\arg \min_{\mathbf{v}} \sum_{\mathbf{x} \in \Omega} \frac{(\mathbf{I}_n(\mathbf{x}) - \mu_I)^2}{2\sigma_I^2} \quad (11.42)$$

En pratique, μ_I peut être approximé par la moyenne de $\mathbf{I}_n$ dans Ω . Les versions continues des critères (11.40) et (11.42) peuvent être linéairement combinées pour définir une énergie spatio-temporelle paramétrique :

$$E_p(\Gamma) = \int_{\Omega} (\mathbf{I}_n(\mathbf{x}) - \mu_I(\Gamma))^2 d\mathbf{x} + \frac{2\sigma_I^2}{\sigma} \int_{\Omega} |\mathbf{I}_n(\mathbf{x}) - \mathbf{I}_{n+1}(\mathbf{x} + \mathbf{v}(\Gamma))| d\mathbf{x} \quad (11.43)$$

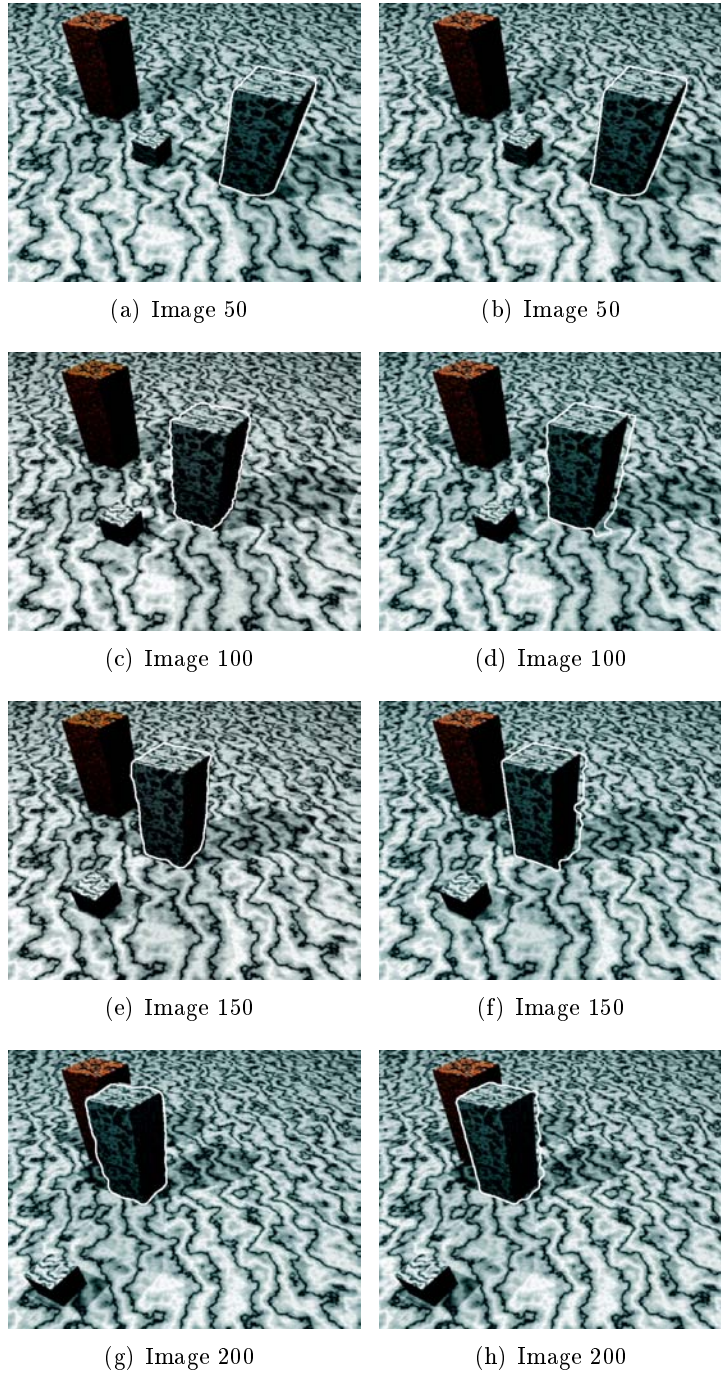


FIG. 11.7 – Segmentation basée mouvement sur la séquence **Marbled blocks** entre les images 50 et 200. La colonne de gauche montre le suivi obtenu avec notre méthode et la colonne de droite le suivi obtenu avec la méthode paramétrique.



FIG. 11.8 – Comparaison d’une approche paramétrique et d’une non-paramétrique sur la séquence **City**.

où

$$\begin{cases} \mu_I(\Gamma) = \frac{\int_{\Omega} \mathbf{I}_n(\mathbf{x}) \, d\mathbf{x}}{\int_{\Omega} d\mathbf{x}} \\ \mathbf{v}(\Gamma) = \arg \min_w \int_{\Omega} |\mathbf{I}_n(\mathbf{x}) - \mathbf{I}_{n+1}(\mathbf{x} + w)| \, d\mathbf{x} \end{cases} \quad (11.44)$$

Pour donner une idée du comportement de chaque terme dans l’énergie paramétrique, la segmentation a été réalisée avec chaque terme indépendamment et en combinant les deux (comme pour l’énergie non-paramétrique, voir section 11.4.2).

Cette approche paramétrique a aussi été testée sur une séquence plus difficile, **Football** (voir figure 11.10).

A la lumière de ces résultats, plusieurs remarques peuvent être faites. Comme cela était attendu, quand des hypothèses paramétriques faites correspondent bien aux distributions réelles (séquence **Flowers and garden** par exemple), l’approche paramétrique permet une segmentation correcte. De plus, quand les distributions paramétriques ne correspondent pas du tout aux distributions réelles lorsque le mouvement est complexe ou bien que l’objet ou le fond est composé de plusieurs couleurs (séquence **Football**), l’approche paramétrique échoue au contraire de l’approche non-paramétrique (voir figure 11.3 et figure 11.4).

11.4.4 Suivi d’objet et estimation du mouvement par morceaux

Dans cette section, un objet d’intérêt est suivi le long de la séquence sur deux séquences de test en utilisant la méthode non-paramétrique proposée. Dans ces deux séquences, l’objet d’intérêt a un mouvement articulé compliqué et est composé de plusieurs couleurs. La séquence **Soccer** est moins complexe que la séquence **Football** parce que cette dernière souffre notamment d’un flou de mouvement (séquence mal échantillonnée). Ces séquences sont appropriées pour comparer l’estimation du mouvement global (sur Ω) et l’approche locale (sur Ω_i). Pour l’estimation de mouvement par morceaux, chaque image a été divisée en 16×16 blocs de 22×18 pixels. Pour rappel, le domaine Ω est définie par $\Omega \cap B_i$. La

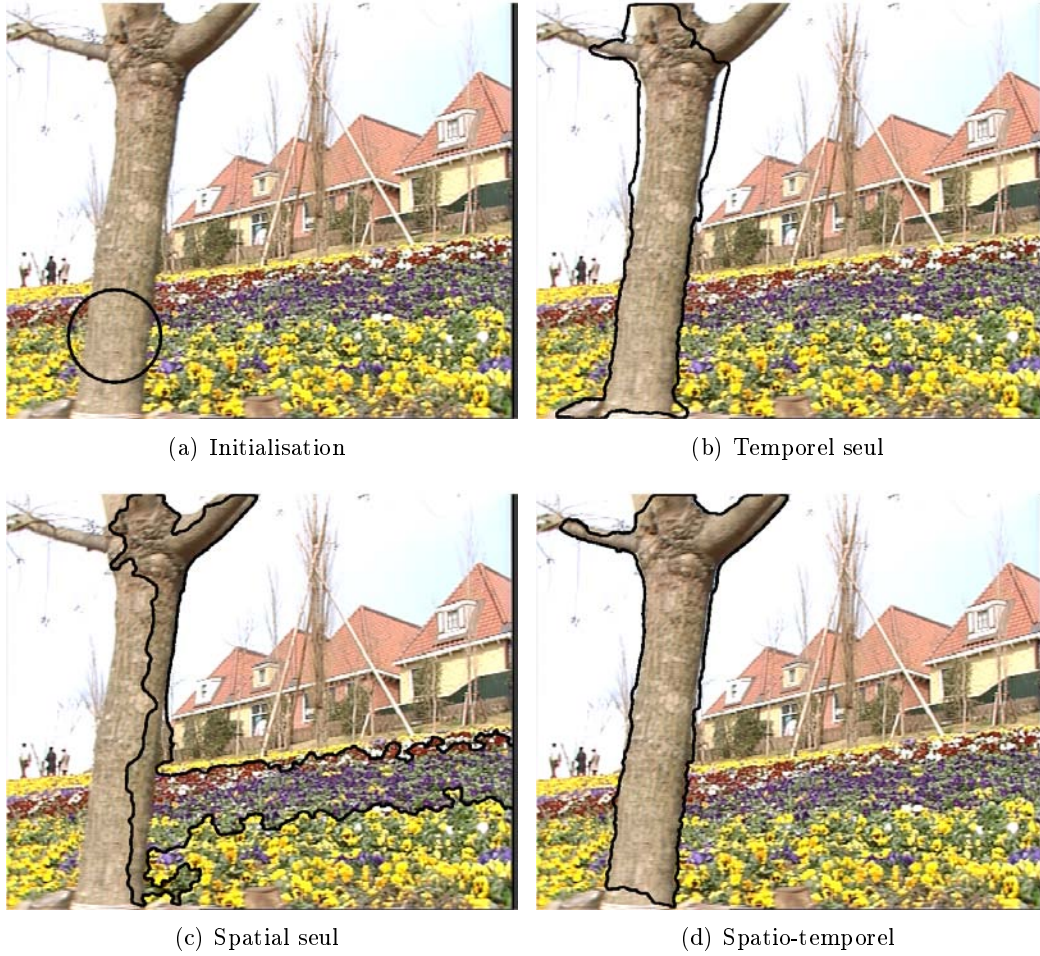


FIG. 11.9 – Segmentation de l'image 237 de la séquence **Flowers and garden** en supposant des distributions paramétriques.

comparaison entre ces deux approches est présentée dans les figures 11.11 et 11.12.

Bien que l'approche locale augmente clairement la qualité de la segmentation, ce n'est pas toujours parfait sur la séquence **Football**. Cela s'explique par la combinaison des effets de flou de mouvement et par le fait que le domaine Ω_i est trop petit (ce qui arrive pour les blocks B_i qui intersectent Γ), ce qui provoque une estimation du mouvement local moins fiable.

11.5 CONCLUSION

Nous avons présenté ici une méthode de segmentation d'objets mobiles avec estimation conjointe du mouvement. Un modèle de mouvement est choisi et l'entropie de l'erreur de prédiction est minimisée afin de segmenter les objets pour lesquels le mouvement est bien compensé, c'est-à-dire les objets qui respectent le modèle de mouvement. Afin d'être plus

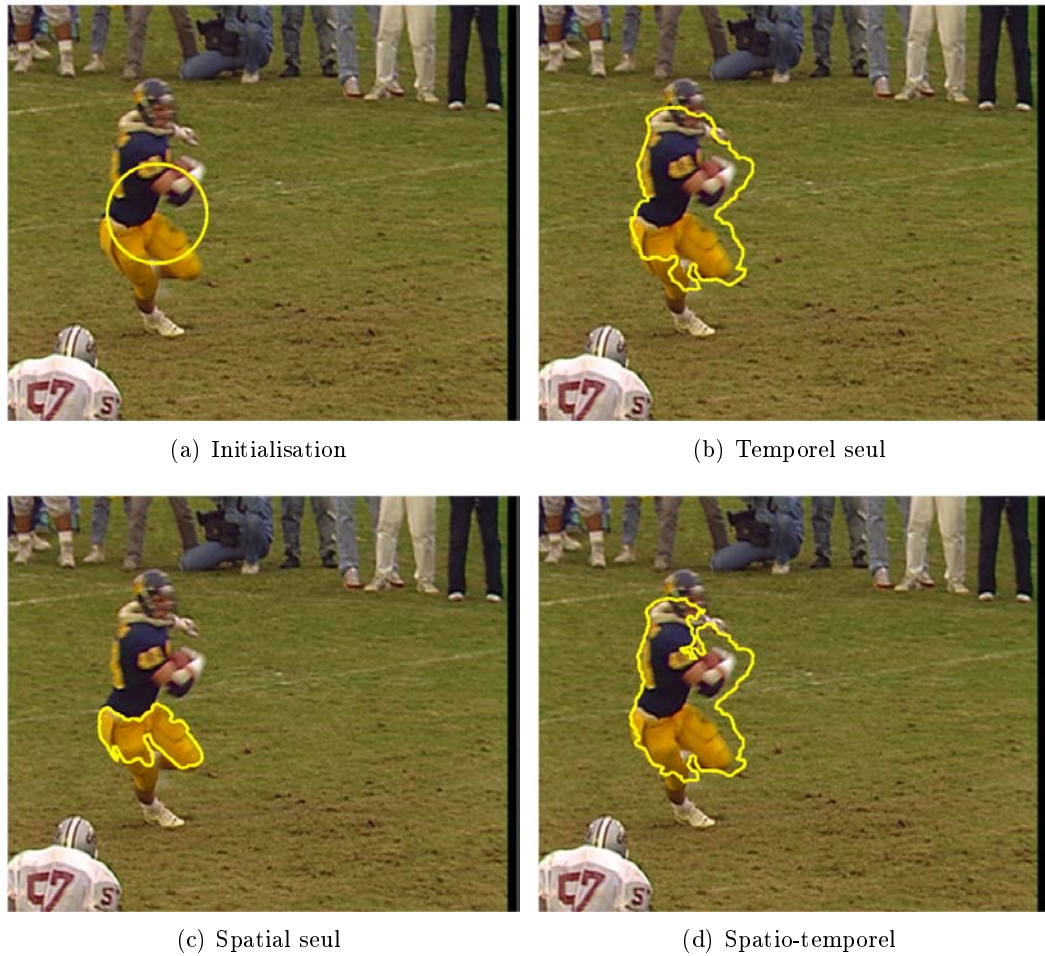


FIG. 11.10 – Segmentation de l'image 72 de la séquence **Football** en supposant des distributions paramétriques.

robuste aux zones homogènes qui peuvent mettre en défaut l'estimateur de mouvement, nous avons introduit de l'information spatiale en minimisant l'entropie jointe de l'erreur de prédiction et de l'intensité lumineuse. De ce fait, notre critère peut segmenter des objets en mouvement même si des zones homogènes sont présentes dans l'objet ou sur le fond. Nous avons appliqué cette méthode à la segmentation d'objets mobiles dans des séquences vidéo. Afin d'être plus robuste quand il s'agit d'objets articulés nous avons introduit un modèle de mouvement local afin de pouvoir segmenter des objets qui se déforment fortement comme par exemple des personnages humains.

En supposant l'indépendance entre les données, nous avons obtenu une fonctionnelle formée d'une somme de plusieurs énergies mais sans paramètre de pondération entre ces énergies. C'est un point intéressant puisque généralement lorsque l'on combine plusieurs énergies dans une fonctionnelle à minimiser, un paramètre de pondération entre les termes est nécessaire et son réglage est délicat.

Des améliorations pourraient être envisagées afin d'obtenir de meilleurs résultats sur des séquences réelles. Nous pourrions tout d'abord imaginer un modèle de mouvement plus proche de la réalité et bien adapté aux objets déformables. Nous pourrions aussi envisager de prendre en compte d'autres caractéristiques, comme par exemple des modèles de texture ou de forme. Cette approche est donc bien adaptée à la prise en compte de plusieurs caractéristiques et une extension à des données provenant de plusieurs modalités est envisageable.



FIG. 11.11 – Suivi d'objet sur la séquence **Soccer** : comparaison entre l'approche du mouvement global (colonne de gauche) et l'approche du mouvement local (colonne de droite) sur les images 164, 172, 182, 192 et 202. La segmentation est plus précise avec l'approche locale (les principales différences sont entourées d'un cercle).



FIG. 11.12 – Suivi d'objet sur la séquence Football : comparaison entre l'approche du mouvement global (colonne de gauche) et l'approche du mouvement local (colonne de droite) sur les images 73, 77, 81, 85 et 89. La segmentation est plus précise avec l'approche locale (les principales différences sont entourées d'un cercle).

Quatrième partie

Conclusion et perspectives

Conclusion

Dans ce mémoire nous nous étions fixés comme objectifs de présenter des méthodes de segmentation d'images et de vidéos en utilisant des distributions non-paramétriques afin de ne pas modéliser les distributions des caractéristiques considérées (intensité lumineuse par exemple). Nous proposons l'utilisation de critères statistiques provenant de la théorie de l'information dans un modèle de contours actifs pour la segmentation d'images et de séquences vidéo.

Nous utilisons une approche variationnelle, c'est-à-dire que le problème de segmentation est formulé comme la minimisation d'un critère. Nous considérons des critères basés régions qui incluent de l'information sur une région entière et pas uniquement sur ses frontières comme cela est le cas avec les critères basés contour. Ces critères peuvent impliquer la moyenne, la variance, le mouvement ou bien d'autres fonctions de leurs caractéristiques. Mais ces fonctions font bien souvent une hypothèse sur la distribution des caractéristiques considérées. Par exemple, considérer une fonction de la moyenne de l'intensité d'une région revient à faire une hypothèse gaussienne sur cette intensité. Dans ce manuscrit nous nous intéressons à ne pas modéliser les densités de probabilité des caractéristiques de l'image, car cette hypothèse n'est pas toujours respectée, par exemple dans une séquence vidéo, à cause des changements d'illumination, des occlusions, des erreurs dans le modèle de mouvement.

Nous nous intéressons donc à des critères non-paramétriques, issus de la théorie de l'information afin de considérer des critères prenant en compte l'information contenue dans l'image. Nous faisons dans nos travaux l'hypothèse d'indépendance de l'intensité lumineuse de pixels voisins. Et dans certains cas, nous supposons aussi l'indépendance des différents canaux des densités de probabilité que nous utilisons.

Nous présentons, les mesures d'information issues de la théorie de l'information qui sont utilisées par la suite pour la segmentation d'images et de séquences vidéo. Nous avons besoin d'estimer les densités de probabilité des données et nous utilisons la méthode de Parzen, une approche non-paramétrique qui utilise un noyau. Nous intégrons ensuite ces distributions dans des critères fonctions de ces distributions. Nous nous intéressons particulièrement à l'entropie qui mesure la quantité d'information contenue dans un signal. Plus le signal est redondant, moins son entropie est élevée. Nous présentons ensuite l'information mutuelle qui mesure la dépendance statistique des deux caractéristiques considérées. Nous introduisons enfin la divergence de Kullback-Leibler qui mesure la dissimilarité entre deux

densités de probabilité. La divergence de Kullback-Leibler est utilisée principalement pour mesurer l'information contenue dans la différence entre une distribution de référence et une distribution théorique ou modélisée.

Nous appliquons ensuite ces critères à la segmentation d'images fixes. Dans le cas de la segmentation d'images en niveaux de gris, nous considérons l'entropie de l'intensité lumineuse. Nous étendons ensuite cette étude en considérant l'entropie jointe des différents canaux couleurs pour le cas de segmentation d'images couleur. Nous appliquons ce critère à la segmentation de visages et à la segmentation d'objets dans des images couleur.

Dans le cas où nous disposons d'une segmentation de référence, nous choisissons de minimiser une distance entre distributions. Nous utilisons donc la divergence de Kullback-Leibler comme mesure de dissimilarité entre les densités de probabilités de la région de référence et de la région courante. Nous appliquons cette mesure à la segmentation de visages dans le cas de séquences vidéo où nous disposons d'une segmentation de référence sur une image et où nous voulons segmenter l'image suivante dans la séquence. Nous pouvons observer qu'à convergence de l'algorithme, les densités de probabilité de l'objet segmenté et du fond sont très proches de celles de référence et que la segmentation est correcte.

Nous considérons enfin un critère de distance à une segmentation de référence, combiné à un critère d'entropie spatiale. La distance est obtenue en appliquant au préalable un algorithme de ligne de partage des eaux sur l'image, ce qui donne une sur-segmentation de l'image. À partir de cette segmentation nous calculons une carte de distances qui sert d'a priori géométrique dans notre critère. Nous combinons ainsi une information radiométrique sur la région et une contrainte géométrique, en forçant le contour actif à se rapprocher au maximum des endroits de fort gradient.

Enfin, nous considérons la segmentation de séquences vidéo et nous utilisons les critères statistiques issus de la théorie de l'information pour segmenter des objets en mouvement. Nous présentons deux façons de considérer le mouvement.

Tout d'abord, nous considérons un champ dense de mouvement, à savoir un vecteur mouvement par pixel. Sur une séquence à caméra fixe, nous calculons directement le flot optique (le champ dense). Sur une séquence à caméra mobile, nous estimons le mouvement de la caméra, et nous compensons ce mouvement pour nous replacer dans un contexte de caméra fixe. Une fois ce champ dense obtenu, nous cherchons à minimiser l'entropie des vecteurs mouvement. Nous cherchons les régions dont les vecteurs mouvement sont les plus homogènes possible.

Nous proposons une seconde approche qui consiste à estimer conjointement le mouvement avec la segmentation. Pour cela nous choisissons un modèle de mouvement pour l'objet d'intérêt et nous minimisons l'entropie de l'erreur de prédiction entre le mouvement estimé et le mouvement réel. Cette méthode segmente correctement pour les séquences vidéo texturées mais sur-segmente les zones homogènes. Pour pallier cela, nous choisissons de prendre en compte la densité de probabilité de l'intensité lumineuse de la région en plus de l'erreur de prédiction. Nous considérons donc l'entropie conjointe de l'erreur de prédiction et de l'intensité lumineuse. En faisant l'hypothèse d'indépendance entre le résiduel et l'intensité lumineuse, nous segmentons correctement des séquences qui comportent des

zones texturées et d'autres plus homogènes.

PERSPECTIVES

Les travaux présentés dans cette thèse mettent en lumière des perspectives qu'il est intéressant de présenter.

Nous avons vu, que pour estimer des densités de probabilité, les méthodes à noyau comme la méthode de Parzen sont adaptées pour des petites dimensions mais que cela n'est plus le cas quand la dimension augmente car elles deviennent alors inefficaces [Sco92]. C'est cette limitation des méthodes à noyau qui nous a poussé à considérer l'indépendance des différents canaux considérés, afin de réduire la dimension du noyau utilisé. Il est donc nécessaire d'envisager d'autres méthodes d'estimation de densités qui soient plus adaptées aux grandes dimensions. Il existe par exemple des méthodes d'estimation basées sur les k -ièmes plus proches voisins qui permettent d'obtenir facilement des densités de probabilité de dimension supérieure [BDB07]. Le principe de cette estimation est de calculer la distance entre chaque point de l'échantillon et ses k plus proches voisins. L'estimateur est construit de façon à adapter la taille des régions au nombre d'observations désiré. Chaque intervalle contient donc k points de l'échantillon. Ce genre de méthode permet d'avoir une largeur de bande adaptée à la densité locale des points de l'échantillon et d'éviter d'obtenir un sous-lissage des données au niveau des queues de distributions. Ce type d'estimation est facilement extensible au cas multivarié en considérant comme distance l'hypersphère centrée en chaque point de l'échantillon. Le choix du paramètre k est moins critique que celui de la largeur de bande pour la méthode de Parzen, en théorie $k = \sqrt{n}$, avec n le nombre de points de l'échantillon. Il est aussi possible d'utiliser des méthodes basées sur les k -ièmes plus proches voisins pour estimer directement l'entropie sans estimer la densité de probabilités.

Il pourrait être envisagé d'utiliser d'autres mesures comme descripteurs. Nous pourrions penser à des mesures comme la divergence de Bregman qui est utilisée en classification [BMDG05] et en géométrie algorithmique [NBN07]. Cette divergence englobe une famille de mesures qui regroupe notamment l'erreur au carré et l'entropie relative.

Combiner de l'information sur l'intensité des pixels à de l'information sur leur localisation permettrait d'obtenir des critères de segmentation pour les textures. Pour cela, nous pourrions utiliser des tenseurs de structure [RBD03] ou bien des modèles de textures non-paramétriques [ATW06].

Nous avons vu que dans le cas où l'on dispose d'une segmentation de référence, une notion de distance entre densité de probabilités peut être intéressante. Il pourrait être envisagé d'étudier d'autres distances que celle choisie dans ce manuscrit. Nous pourrions penser à la divergence de Jensen-Shannon, la distance de Kolmogorov, la distance de Bhattacharyya ou bien à des distances entre fonctions de répartition.

Afin d'apporter un *a priori* sur le type d'objet que l'on souhaite segmenter, nous pourrions aussi faire intervenir un *a priori* de forme afin de segmenter des objets qui ont une forme connue à l'avance. De nombreux travaux ont déjà été effectués [LGF00, RP02, CTH⁺02, TYW⁺03, CS03, CFK05] à ce sujet et il serait pertinent de combiner des notions de forme et de photogrammétrie.

Cinquième partie

Annexes

Annexe A

Distance entre fonctions de répartition

Dans cette annexe, nous présentons un critère basé sur une distance entre fonctions de répartition. Nous dérivons ce critère pour parvenir à l'équation d'évolution du contour actif en utilisant les outils de dérivation de forme.

A.1 DÉFINITION DU CRITÈRE

Etant donné une portion Ω de l'image $\mathbf{I}$ et $\mathbf{x}$ un pixel de la région Ω , la densité de probabilité de I dans la région Ω est estimée avec la méthode de Parzen :

$$p(\alpha, \Omega) = \frac{1}{|\Omega|} \int_{\Omega} K_h(\mathbf{I}(\mathbf{x}) - \alpha) d\mathbf{x} \quad (\text{A.1})$$

avec K_h la densité-noyau de l'estimation de Parzen.

La fonction de répartition de $\mathbf{I}$ est alors définie par :

$$H(\lambda, \Omega) = \int_{-\infty}^{\lambda} p(\alpha, \Omega) d\alpha \quad (\text{A.2})$$

Notons Ω_{ref} une région de référence. Nous définissons la distance D entre la fonction de répartition de la région de référence Ω_{ref} , notée $H(\lambda, \Omega_{ref})$ et la fonction de répartition de la région Ω , notée $H(\lambda, \Omega)$:

$$D(\Omega) = \int_{\Lambda} \varphi(H(\lambda, \Omega), H(\lambda, \Omega_{ref})) d\lambda \quad (\text{A.3})$$

où Λ est l'ensemble des valeurs de $\mathbf{I}(\mathbf{x})$. Cette distance pourra être utilisée lorsque l'on dispose par exemple d'une segmentation de référence et que l'on cherche une région dont les caractéristiques se rapprochent de celles de la région de référence. Minimiser cette distance reviendra à minimiser la différence entre les distributions et par conséquent, la région obtenue aura une distribution proche de celle de référence.

Afin d'utiliser ce critère dans un schéma de contour actif, nous devons le dériver pour parvenir à l'équation d'évolution du contour actif.

A.2 DÉRIVATION DU CRITÈRE PAR LES GRADIENTS DE FORME

Nous utilisons les outils de dérivation des gradients de forme pour dériver le critère (A.3).

La dérivée eulérienne de la densité de probabilité p dans la direction $\mathbf{V}$ est donnée par :

$$p'(\alpha, \Omega, \mathbf{V}) = \frac{1}{|\Omega|} \int_{\Gamma} (p(\alpha, \Omega) - K_h(\mathbf{I}(\mathbf{x}) - \alpha))(\mathbf{V} \cdot \mathbf{N}) d\mathbf{x} \quad (\text{A.4})$$

Alors la dérivée de H est :

$$H'(\lambda, \Omega, \mathbf{V}) = \int_{-\infty}^{\lambda} p'(\alpha, \Omega, \mathbf{V}) d\alpha \quad (\text{A.5})$$

$$= \int_{-\infty}^{\lambda} \frac{1}{|\Omega|} \int_{\Gamma} (p(\alpha, \Omega) - K_h(\mathbf{I}(\mathbf{x}) - \alpha))(\mathbf{V} \cdot \mathbf{N}) d\mathbf{x} d\alpha \quad (\text{A.6})$$

$$= \frac{1}{|\Omega|} \int_{\Gamma} \left(\int_{-\infty}^{\lambda} p(\alpha, \Omega) d\alpha - \int_{-\infty}^{\lambda} K_h(\mathbf{I}(\mathbf{x}) - \alpha) d\alpha \right) (\mathbf{V} \cdot \mathbf{N}) d\mathbf{x} \quad (\text{A.7})$$

$$= \frac{1}{|\Omega|} \int_{\Gamma} \left(H(\lambda, \Omega) - \int_{-\infty}^{\lambda} K_h(\mathbf{I}(\mathbf{x}) - \alpha) d\alpha \right) (\mathbf{V} \cdot \mathbf{N}) d\mathbf{x} \quad (\text{A.8})$$

Nous pouvons alors calculer la dérivée du critère (A.3) :

$$D'(\Omega, \mathbf{V}) = \int_{\Lambda} H'(\lambda, \Omega, \mathbf{V}) \varphi'(H(\lambda, \Omega), H(\lambda, \Omega_{ref})) d\lambda \quad (\text{A.9})$$

$$= \int_{\Lambda} \frac{1}{|\Omega|} \left[\int_{\Gamma} \left(H(\lambda, \Omega) - \int_{-\infty}^{\lambda} K_h(\mathbf{I}(\mathbf{x}) - \alpha) d\alpha \right) (\mathbf{V} \cdot \mathbf{N}) d\mathbf{x} \right] \varphi'(H(\lambda, \Omega), H(\lambda, \Omega_{ref})) d\lambda \quad (\text{A.10})$$

Nous changeons l'ordre d'intégration :

$$D'(\Omega, \mathbf{V}) = \frac{1}{|\Omega|} \int_{\Gamma} \int_{\Lambda} \left[H(\lambda, \Omega) - \int_{-\infty}^{\lambda} K_h(\mathbf{I}(\mathbf{x}) - \alpha) d\alpha \right] \varphi'(H(\lambda, \Omega), H(\lambda, \Omega_{ref})) d\lambda (\mathbf{V} \cdot \mathbf{N}) d\mathbf{x} \quad (\text{A.11})$$

De cette dérivée nous pouvons en déduire l'équation d'évolution :

$$\frac{\partial \Gamma}{\partial \tau} = -\frac{1}{|\Omega|} \int_{\Lambda} \left[H(\lambda, \Omega) - \int_{-\infty}^{\lambda} K_h(\mathbf{I}(\mathbf{x}) - \alpha) d\alpha \right] \varphi'(H(\lambda, \Omega), H(\lambda, \Omega_{ref})) d\lambda \mathbf{N} \quad (\text{A.12})$$

$$= -\frac{1}{|\Omega|} \int_{\Lambda} H(\lambda, \Omega) \varphi'(H(\lambda, \Omega), H(\lambda, \Omega_{ref})) d\lambda \mathbf{N} \quad (\text{A.13})$$

$$+ \frac{1}{|\Omega|} \int_{\Lambda} \int_{-\infty}^{\lambda} K_h(\mathbf{I}(\mathbf{x}) - \alpha) d\alpha \cdot \varphi'(H(\lambda, \Omega), H(\lambda, \Omega_{ref})) d\lambda \mathbf{N}$$

Cette équation d'évolution permet de faire évoluer le contour actif vers une région qui aura des propriétés statistiques proches de celles de référence.

Méthode de segmentation par *watersheds*, l'algorithme de la ligne de partage des eaux

B.1 DÉFINITION DE L'ALGORITHME DE LA LIGNE DES PARTAGE DES EAUX

L'algorithme de la ligne de partage des eaux (LPE) est un algorithme issu du domaine de la morphologie mathématique qui consiste en la segmentation d'une image en régions homogènes. Il a été introduit par Digabel et Lantuéjoul [DL78], puis amélioré par Beucher et Lantuéjoul [BL79]. Plusieurs implémentations ont permis de l'utiliser pour la segmentation d'images [BM93, VS91]. Pour une revue des méthodes basées sur les watersheds, voir [RM01].

Le principe de cet algorithme est de considérer l'image en termes géographiques. L'image est ainsi perçue en relief si à chaque niveau de gris est associé une altitude. La ligne de partage des eaux est alors la crête formant la limite entre deux bassins versants, un bassin versant étant la zone associée à un minimum régional telle qu'une goutte d'eau tombant dans cette zone et suivant la ligne de plus grande pente s'arrêtera dans ce minimum (voir figure B.1).

La transposition de la définition intuitive de la ligne de partage des eaux ne se formalise pas simplement en termes mathématiques. Souvent la LPE est définie d'après l'algorithme qui permet de la construire. Un des algorithmes les plus populaires pour l'implémenter est l'algorithme d'immersion qui consiste à remplir progressivement les bassins versants à partir des minima, pour déterminer leurs limites [VS91]. Une autre implémentation est basée sur l'utilisation de fonctions de distances géodésiques [Mey88].

La LPE est un algorithme qui fournit une partition de l'image et qui est appliqué non pas à l'image directement mais à son gradient morphologique. Cela crée donc des lignes de partage des eaux aux points de discontinuités du niveau de gris.

Les lignes de partage des eaux délimitent en théorie des régions homogènes en niveaux de gris. En pratique, la LPE produit une importante sur-segmentation de l'image, à cause du bruit dans l'image et des irrégularités locales du gradient. Pour limiter cela, des méthodes utilisant des marqueurs ont été développées [BM93]. Le principe consiste à inonder

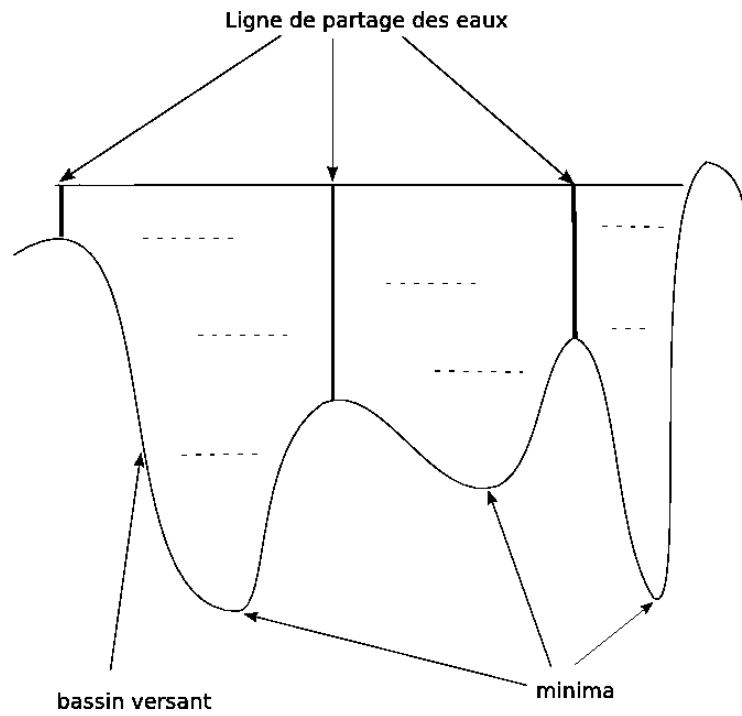


FIG. B.1 – Ligne de partage des eaux

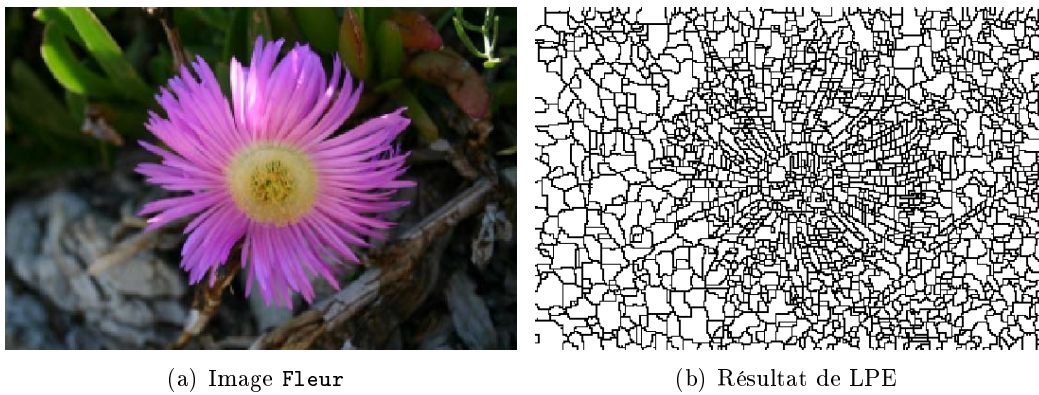


FIG. B.2 – Image et résultat de l'algorithme de la ligne de partage des eaux.

la surface topographique à partir de marqueurs prédéfinis. Des méthodes utilisent un filtrage préalable du gradient de l'image tandis que d'autres méthodes adoptent une approche hiérarchique qui mélange les deux approches précédentes sur différents niveaux en utilisant des graphes.

La figure B.2 montre le résultat de l'algorithme de LPE sur l'image **Fleur**.

B.2 DIMINUTION DU NOMBRE DE RÉGIONS

A cause de la sur-segmentation obtenue en utilisant simplement la LPE, nous ne pouvons créer la carte de distances directement. Nous allons diminuer le nombre de régions existant dans cette sur-segmentation en utilisant un arbre de partition binaire [SG00]. Pour limiter le nombre de régions résultant, nous effectuons au préalable un filtrage du gradient de l'image en utilisant des filtres issus de la morphologie mathématique qui ont la propriété de préserver les contours de l'image. Ensuite nous utilisons un arbre de partition binaire pour représenter les régions obtenues avec la LPE. L'arbre binaire sera créé en calculant un critère de similarité entre les régions. Les deux régions voisines les plus similaires d'après ce critère seront ensuite fusionnées et formeront une nouvelle région. Pas à pas, l'arbre de partition binaire sera ainsi créé. Les régions initiales issues de la sur-segmentation seront les feuilles de l'arbre. Quand deux régions sont fusionnées, la nouvelle région créée est appelé noeud parent dans l'arbre et les deux régions fusionnées sont appelées noeuds enfants. Les régions enfants dans la partie basse de l'arbre seront généralement très similaires tandis que les régions enfants dans le haut de l'arbre seront assez différentes. Le nombre de régions final sera déterminé par la hauteur dans l'arbre binaire.

Le critère utilisé pour calculer la disparité entre deux régions est un critère spatio-temporel composé d'un terme spatial et d'un terme temporel. Dans le cas d'images fixes, le terme temporel est nul. Le terme spatial mesure l'intensité moyenne de la différence entre deux régions adjacentes. Le terme temporel mesure la différence moyenne entre deux régions de la différence d'intensité moyenne d'une région entre deux images successives dans une séquence vidéo. Une région qui se déplace ne sera ainsi pas fusionnée avec une région immobile, et réciproquement. Le critère spatio-temporel mesure donc la similarité de deux régions, plus ce critère est fort plus la disparité entre les régions est forte. La figure B.3 montre la partition initiale d'une image issue de la séquence **Mother and daughter** en 15 régions et l'arbre de partition binaire correspondant.

La figure B.4 présente les résultats de fusion sur une image de la séquence **Mother and daughter** selon le nombre de régions final souhaité.

B.3 UTILISATION DE CETTE SEGMENTATION POUR LA MÉTHODE DES CONTOURS ACTIFS

Nous utilisons cette segmentation par ligne de partage des eaux afin de calculer un *a priori* de distance dans le chapitre 8.

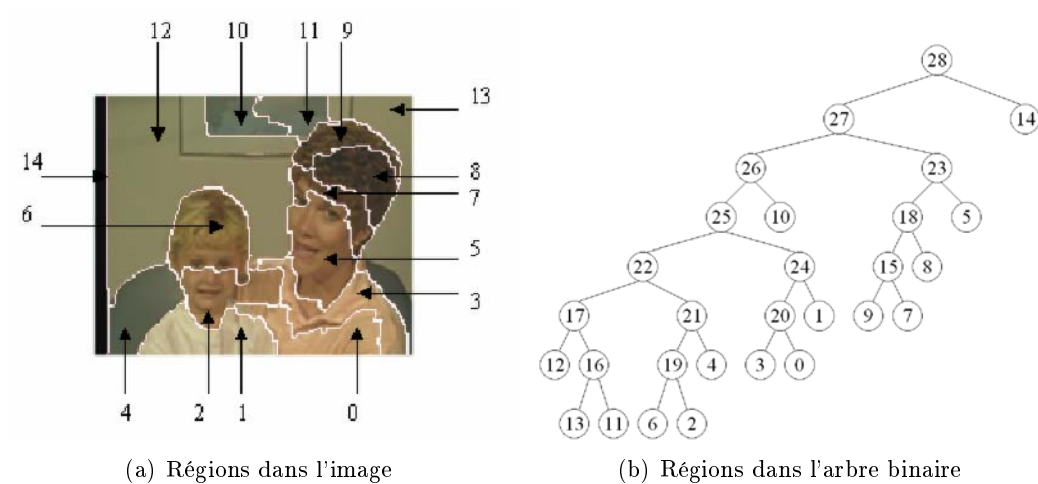


FIG. B.3 – Illustration de l'arbre de partition binaire.

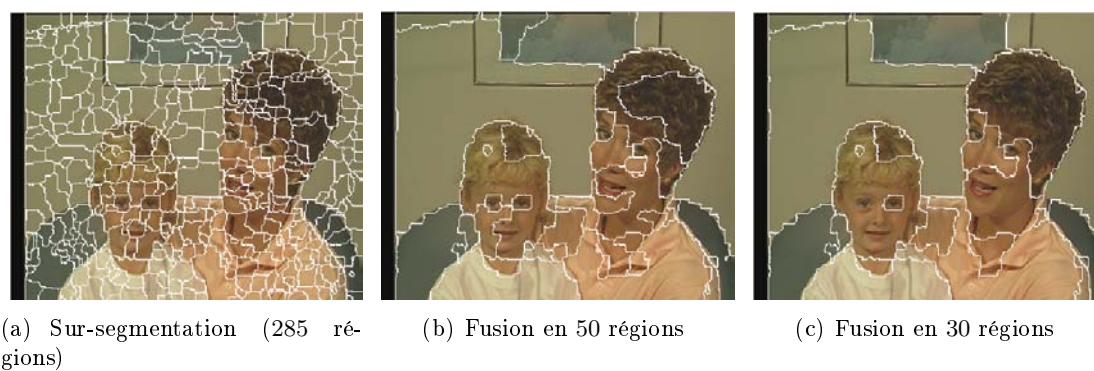


FIG. B.4 – Fusion de la sur-segmentation.

Compensation de la caméra pour des séquences vidéo à caméra mobile

Le principe de la compensation de caméra pour les séquences vidéo à caméra mobile consiste à estimer le mouvement de la caméra et ensuite à le compenser afin de se replacer dans un contexte de caméra fixe.

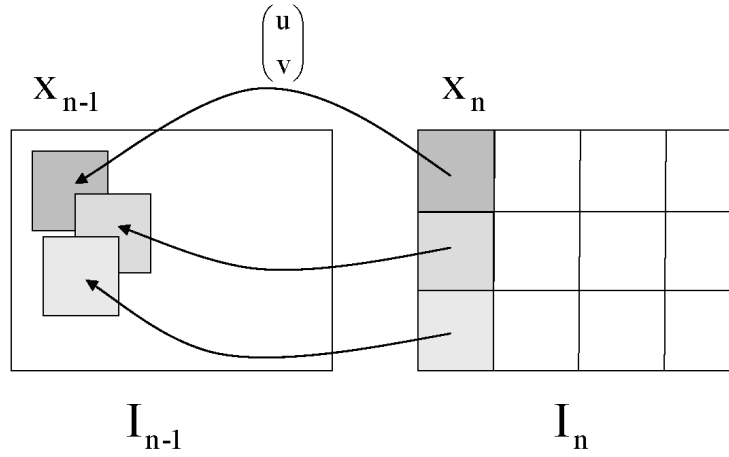
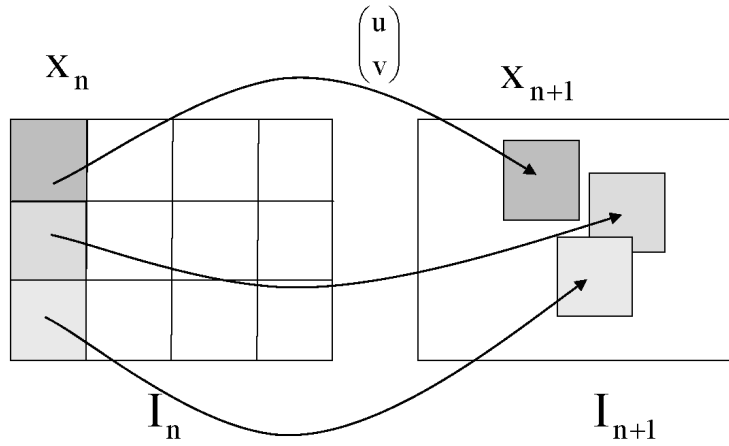
C.1 ESTIMATION DU MOUVEMENT DE LA CAMÉRA

Le mouvement de la caméra est le mouvement dominant de l'image et les objets en mouvement sont les régions ayant un mouvement différent du mouvement dominant. Il s'agit donc de déterminer le mouvement global de l'image en déterminant choisissant un modèle de mouvement et en déterminant ses paramètres. Nous choisissons un modèle affine à 6 paramètres qui est un bon compromis entre précision et coût calculatoire [OB95, JbBA02]. Le principe de l'estimation du mouvement de la caméra se décompose en deux étapes : tout d'abord il s'agit de découper l'image en blocs et d'appliquer un algorithme de mise en correspondance de ces blocs (méthode de *block matching*) afin d'associer un vecteur mouvement par blocs, puis de définir le modèle de mouvement le plus représentatif pour l'ensemble des vecteurs mouvement par une méthode de minimisation. Nous cherchons à estimer le mouvement de la caméra entre deux images successives.

C.1.1 Mise en correspondance de blocs ou *block matching*

Le principe de cette méthode consiste à découper l'image en blocs (de taille 8x8 pixels ou 16x16 pixels généralement) et à chercher pour chaque bloc le bloc de l'autre image qui lui ressemble le plus. Le déplacement de ce bloc est calculé comme étant la différence d'emplacement entre les deux blocs et noté $\mathbf{v}_n = (u, v)^T$. Cette recherche de mise en correspondance peut se faire en cherchant le bloc dans l'image précédente (méthode *backward*) ou dans l'image suivante (méthode *forward*) (voir figure C.1 et figure C.2).

Il existe plusieurs méthodes pour chercher le bloc correspondant, de la méthode exhaustive à la méthode à trois étapes en passant par la recherche en diamant. Chacune de ces méthodes cherche le bloc correspondant dans une fenêtre de recherche autour de la position du bloc que l'on cherche à apparier. C'est la façon de rechercher le bloc dans cette fenêtre de recherche qui varie selon les méthodes. La méthode exhaustive cherche toutes les

FIG. C.1 – Estimation du mouvement par méthode *backward*.FIG. C.2 – Estimation du mouvement par méthode *forward*.

positions possible dans la fenêtre de recherche. La méthode en trois étapes cherche d'abord le bloc dans un voisinage à 4 pixels de distance, puis à 2 pixels puis à 1 pixel. La méthode en diamant utilise un motif de recherche en diamant pour effectuer cette même recherche. Toujours est-il que ces deux dernière méthodes, dites rapides, donnent de bons résultats avec un coût calculatoire faible.

Une fois cette mise en correspondance obtenue pour chaque bloc de l'image par méthode *backward*, nous devons estimer les paramètres du mouvement affine qui représente le mieux le mouvement entre les deux images.

C.1.2 Estimation robuste du modèle de mouvement de la caméra

Le mouvement apparent $\mathbf{v}_n$ d'un point $\mathbf{x}_n = (x, y)^T$ entre deux images I_{n-1} et I_n est donné par :

$$\mathbf{w}_n(\mathbf{x}_n) = A_n \mathbf{x}_n + t_n = \begin{pmatrix} a_{11}^n & a_{12}^n \\ a_{21}^n & a_{22}^n \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_1^n \\ t_2^n \end{pmatrix}. \quad (\text{C.1})$$

où A_n et t_n représentent les paramètres du modèle affine de mouvement.

Nous cherchons à minimiser la différence entre $A_n \mathbf{x}_n + t_n$ l'image du pixel $\mathbf{x}_n$ par le modèle de mouvement de la caméra, et $\mathbf{v}_n(\mathbf{x}_n)$ son correspondant estimé par *block-matching*. Nous choisissons de minimiser le critère suivant :

$$G(A_n, t_n) = \sum_{\mathbf{x}_n \in \Omega_I} \varphi(|\mathbf{v}_n(\mathbf{x}_n) - A_n \mathbf{x}_n - t_n|). \quad (\text{C.2})$$

La fonction φ permet d'éliminer les points aberrants dus principalement au mouvement des objets en mouvement. Nous choisissons pour φ l'estimateur de Geman et McLure [GML85] :

$$\varphi(r) = \frac{r^2}{1 + r^2}$$

Pour minimiser le critère (C.2), nous utilisons le théorème semi-quadratique énoncé par Charbonnier *et al.* [CBFAB97] et le problème de minimisation peut être formulé de façon équivalente comme :

$$(A_n, t_n) = \arg \min_{(A_n, t_n)} \sum_{\Omega_I} w_r r^2 \quad (\text{C.3})$$

où $r = |u_n - A_n \mathbf{x} - t_n|$ et $w_r = \frac{\varphi'(r)}{2r}$.

La minimisation est effectuée en utilisant l'algorithme ARTUR [CBFAB97] basé sur les minimisations alternées :

Initialisation :

$$k = 0, \quad w_r^0 = 1, \quad r^0 = |\mathbf{v}_n - A_n^0 \mathbf{x}_n - t_n^0|$$

$$(A_n^0, t_n^0) = \arg \min_{(A_n, t_n)} \sum_{\Omega_i} (r^0)^2$$

Répéter :

$$r^k = |\mathbf{v}_n - A_n^k \mathbf{x}_n - t_n^k|$$

$$w_r^k = \frac{\varphi'(r^k)}{2r^k}$$

$$(A_n^{k+1}, t_n^{k+1}) = \arg \min_{(A_n, t_n)} \sum_{\Omega_i} w_r^k (r^k)^2$$

jusqu'à convergence.

La minimisation de $\sum_{\Omega_i} w_r^k (r^k)^2$ est effectuée en utilisant une méthode de descente de gradient. Nous obtenons ainsi les six paramètres du modèle de mouvement.

C.2 COMPENSATION DU MOUVEMENT DE LA CAMÉRA

Le principe de la compensation de caméra consiste à mettre pour chaque pixel de l'image compensée I_{n-1}^{comp} l'intensité du pixel qui lui correspond dans l'image I_{n-1} :

$$I_{n-1}^{comp}(\mathbf{x}_n) = I_{n-1}(A_n \mathbf{x}_n + t_n) \quad (\text{C.4})$$

Les pixels manquants sont calculés par interpolation classique.

En estimant un modèle affine à 6 paramètres entre deux images successives, I_{n-1} et I_n et en appliquant la transformation correspondante à l'image I_{n-1} , nous créons une image compensée de telle façon que le fond reste relativement fixe entre I_{n-1}^{comp} et I_n . La figure C.3 illustre ce phénomène. Le niveau de gris représente la différence de l'intensité lumineuse.

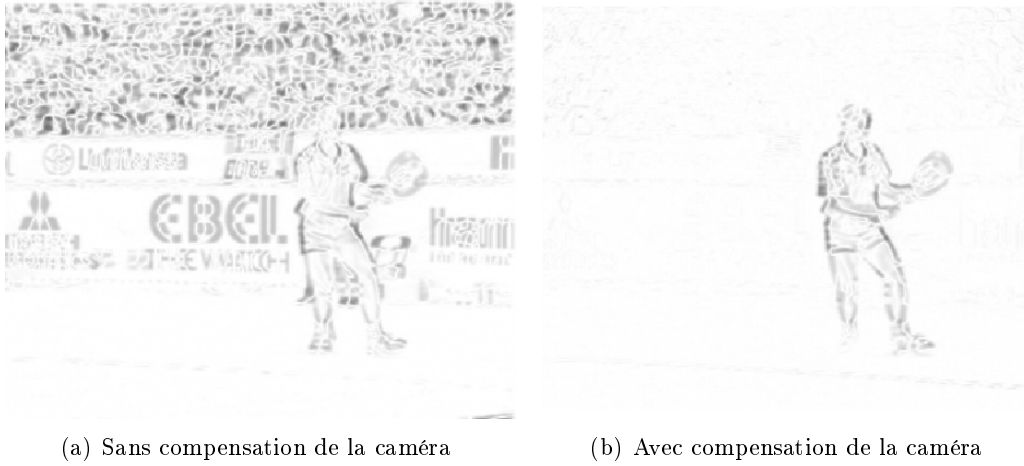


FIG. C.3 – Différence de l'intensité lumineuse entre I_{n-1} et I_n sans compensation de la caméra (à gauche) et avec compensation (à droite).

La figure C.4 montre l'effet de la compensation de caméra dans le contexte de l'estimation du flot optique. Sans compensation de caméra il semble impossible de segmenter le joueur de tennis. Cela est dû au fait que le mouvement de zones homogènes comme le terrain de tennis ne sont pas estimées correctement.

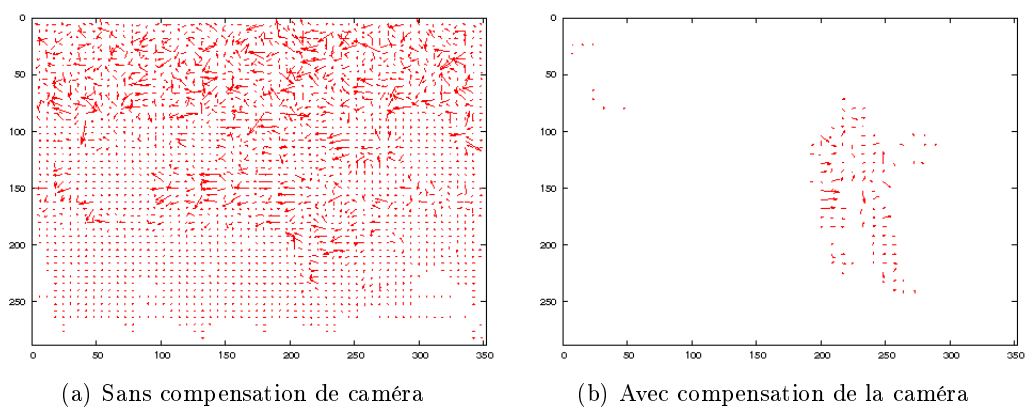


FIG. C.4 – Flot optique sans (à gauche) et avec (à droite) compensation de la caméra.

Bibliographie

- [AABG07] T. ANDRÉ, M. ANTONINI, M. BARLAUD, et R.M. GRAY. Entropy-based bit allocation for wavelet image compression. *IEEE Transactions on Image Processing*, à paraître, 2007. [41](#)
- [AAGLMA⁺00] C. ATAE-ALLAH, J. F. GOMEZ-LOPERA, J. MARTINEZ-AROZA, R. ROMAN-ROLDAN, et P. LUQUE-ESCAMILLA. Image segmentation by jensen-shannon divergence. application to measurement of interfacial tension. Dans *International Conference on Pattern Recognition*, volume 03, pages 3383–3386, Los Alamitos, CA, USA, 2000. IEEE Computer Society. [80](#)
- [ABFJB03] G. AUBERT, M. BARLAUD, O. FAUGERAS, et S. JEHAN-BESSON. Image segmentation using active contours : Calculus of variations or shape gradients ? *SIAM Applied Mathematics*, 1(2) :2128–2145, 2003. [17](#), [47](#)
- [AL76] I.A. AHMAD et P.E. LIN. A nonparametric estimation of the entropy for absolutely continuous distributions. *IEEE Transactions on Information Theory*, 22 :372–375, 1976. [40](#)
- [AS95] D. ADALSTEINSSON et J.A. SETHIAN. A fast level set method for propagating interfaces. *Journal of Computational Physics*, 118 :269 – 277, 1995. [28](#)
- [ATW06] S. P. AWATE, T. TASDIZEN, et R. T. WHITAKER. Unsupervised texture segmentation with nonparametric neighborhood statistics. Dans *European Conference on Computer Vision (2)*, pages 494–507, 2006. [145](#)
- [AWS99] L. ALVAREZ, J. WEICKERT, et J. SÁNCHEZ. A scale-space approach to nonlocal optical flow calculations. Dans *International Conference on Scale-Space Theories in Computer Vision*, pages 235–246, 1999. [100](#)
- [BA96] M. J. BLACK et P. ANANDAN. The robust estimation of multiple motions : parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1) :75–104, 1996. [118](#)
- [BD94] A. BERLINET et L. DEVROYE. A comparison of kernel density estimates. *Publications de l'Institut de Statistique de l'Université de Paris*, 38(3) :3–59, 1994. [33](#), [34](#)

- [BDB07] S. BOLTZ, E. DEBREUVE, et M. BARLAUD. High-dimensional statistical distance for region-of-interest tracking : Application to combining a soft geometric constraint with radiometry. Dans *IEEE International Conference on Computer Vision and Pattern Recognition*, Minneapolis, USA, 2007. [123](#), [145](#)
- [BDGvdM97] J. BEIRLANT, E. DUDEWICZ, L. GYORFI, et E. van der MEULEN. Nonparametric entropy estimation : An overview. *International Journal of the Mathematical Statistics Sciences*, 6 :17–39, 1997. [40](#)
- [Bes86] J. BESAG. On the statistical analysis of dirty pictures (with discussion). *Journal of Royal Statistical Society, Series B*, 48 :259–302, 1986. [47](#)
- [BF93] P. BOUTHEMY et E. FRANCOIS. Motion segmentation and qualitative dynamic scene analysis from an image sequence. *International Journal of Computer Vision*, 10(2) :157–182, avril 1993. [101](#)
- [BJ01] Y. BOYKOV et M.-P. JOLLY. Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images. Dans *International Conference on Computer Vision*, pages 105–112, 2001. [17](#)
- [BL79] S. BEUCHER et C. LANTUEJOUL. Use of watersheds in contour detection. Dans *Proc. of the International Workshop on Image Processing, Real-Time Edge and Motion Detection /Estimation*, pages 17–21, Rennes, septembre 1979. [151](#)
- [BM93] S. BEUCHER et F MEYER. *Mathematical Morphology in Image Processing*, chapitre The morphological approach to segmentation : the watershed transformation, pages 433–481. Marcel Dekker, 1993. [4](#), [151](#)
- [BMDG05] A. BANERJEE, S. MERUGU, I. S. DHILLON, et J. GHOSH. Clustering with bregman divergences. *Journal of Machine Learning Research*, 6 :1705–1749, 2005. [145](#)
- [BRB⁺04] A. BLAKE, C. ROTHER, M. BROWN, P. PÉREZ, et P. H. S. TORR. Interactive image segmentation using an adaptive gmmrf model. Dans *European Conference on Computer Vision (1)*, pages 428–441, 2004. [17](#)
- [BRDW03] T. BROX, M. ROUSSON, R. DERICHE, et J. WEICKERT. Unsupervised segmentation incorporating colour, texture, and motion. Dans N. PETKOV et M. A. WESTENBERG, éditeurs, *Computer Analysis of Images and Patterns*, volume 2756 de *LNCS*, pages 353–360, Groningen, The Netherlands, août 2003. Springer. [101](#), [121](#)
- [Bro66] P. BRODATZ. *Textures : A Photographic Album for Arts and Design*. Dover Publishing Co., 1966. [53](#)
- [BW04] T. BROX et J. WEICKERT. Level set based segmentation of multiple objects. Dans C.E. RASMUSSEN, H.H. BÜLTHOFF, M.A. GIESE, et B. SCHÖLKOPF, éditeurs, *Pattern Recognition (Proc. DAGM)*, volume 3175 de *LNCS*, pages 415–423, Tübingen, Germany, août 2004. Springer. [47](#)
- [BWS02] A. BRUHN, J. WEICKERT, et C. SCHNÖRR. Combining the advantages of local and global optic flow methods. Dans *Proceedings of the 24th DAGM*

- Symposium on Pattern Recognition*, pages 454–462. Springer-Verlag, 2002. [100](#)
- [CBA93] L. COHEN, E. BARDINET, et N. AYACHE. Surface reconstruction using active contour models. Dans *SPIE Conference on Geometric Methods in Computer Vision*, San Diego, CA, 1993. [16](#), [28](#)
- [CBFAB97] P. CHARBONNIER, L. BLANC-FÉRAUD, G. AUBERT, et M. BARLAUD. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6(2) :298–311, février 1997. [100](#), [157](#)
- [CCCD93] V. CASELLES, F. CATTE, T. COLL, et F. DIBOS. A geometric model for active contours. *Numerische Mathematik*, 66 :1–31, 1993. [15](#)
- [CEN07] T. CHAN, S. ESEDOGLU, et K. NI. Histogram based segmentation using wasserstein distances. Dans F. SGALLARI, A. MURLI, et N. PARAGIOS, éditeurs, *Proc. International Conference on Scale Space and Variational Methods in Computer Vision*, volume 4485 de *LNCS*, pages 697–708, Ischia, Italy, mai 2007. Springer. [80](#)
- [CFK05] G. CHARPIAT, O. FAUGERAS, et R. KERIVEN. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Foundations of Computational Mathematics*, 5(1) :1–58, 2005. [146](#)
- [Cho91] D. CHOPP. Computing minimal surfaces via level set curvature flows. Rapport technique, Lawrence Berkeley Laboratory, University of Berkeley, CA, 1991. [28](#)
- [CKS97] V. CASELLES, R. KIMMEL, et G. SAPIRO. Geodesic active contours. *International Journal of Computer Vision*, 22(1) :61–79, 1997. [14](#), [22](#)
- [CM99] D. COMANICIU et P. MEER. Mean shift analysis and applications. Dans *International Conference on Computer Vision*, volume 2, pages 1197–1203, 1999. [119](#)
- [Coh91] L. COHEN. On active contour models and balloons. *Computer Vision, Graphics and Image Processing : Image Understanding*, 53 :211–218, mars 1991. [14](#), [15](#), [28](#)
- [Com03] D. COMANICIU. Nonparametric information fusion for motion estimation. Dans *Computer Vision and Pattern Recognition*, volume 1, pages 59–68, 2003. [119](#)
- [COS06] D. CREMERS, S. J. OSHER, et S. SOATTO. Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. *International Journal of Computer Vision*, 69(3) :335–351, septembre 2006. [47](#)
- [CRB99] C. CHESNAUD, P. RÉFRÉGIER, et V. BOULET. Statistical region snake-based segmentation adapted to different physical noise models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21 :1145–1156, novembre 1999. [16](#), [17](#)
- [CRD07] D. CREMERS, M. ROUSSON, et R. DERICHE. A review of statistical approaches to level set segmentation : Integrating color, texture, motion and

- shape. *International Journal of Computer Vision*, 72(2) :195–215, 2007. [47](#)
- [Cre06] D. CREMERS. Dynamical statistical shape priors for level set based tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(8) :1262–1273, août 2006. [47](#)
- [CS03] D. CREMERS et C. SCHNÖRR. Statistical shape knowledge in variational motion segmentation. *Image and Vision Computing*, 21(1) :77–86, 2003. [101](#), [146](#)
- [CS05] D. CREMERS et S. SOATTO. Motion competition : A variational framework for piecewise parametric motion segmentation. *International Journal of Computer Vision*, 62(3) :249–265, mai 2005. [101](#)
- [CT91] T. M. COVER et J. A. THOMAS. *Elements of information theory*. Wiley-Interscience, New York, NY, USA, 1991. [39](#), [43](#)
- [CTH⁺02] Y. CHEN, S. TAGARE, H.D.and Thiruvankadam, F. HUANG, D. WILSON, K.S. GOPINATH, R.W. BRIGGS, et E.A. GEISER. Using prior shapes in geometric active contours in a variational framework. *International Journal of Computer Vision*, 50(3) :315–328, 2002. [146](#)
- [CV01] T. CHAN et L. VESE. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2) :266–277, 2001. [17](#), [32](#), [47](#)
- [DBAD01] E. DEBREUVE, M. BARLAUD, G. AUBERT, et J. DARCOURT. Space time segmentation using level set active contours applied to myocardial gated SPECT. *IEEE Transactions on Medical Imaging*, 20(7) :643–659, juillet 2001. [17](#)
- [Dev94] L.P. DEVROYE. On good deterministic smoothing sequences for kernel density estimates. *Annals of Statistics*, 22(2) :886–889, 1994. [33](#)
- [DGBA05] E. DEBREUVE, M. GASTAUD, M. BARLAUD, et G. AUBERT. A region-based joint motion computation and segmentation on a set of frames. Dans *Workshop on Image Analysis for Multimedia Interactive Services*, Montreux, Suisse, avril 13-15 2005. [101](#)
- [DHS00] R. DUDA, P. HART, et D. STORK. *Pattern Classification (2nd Edition)*. Wiley-Interscience, novembre 2000. [32](#)
- [DL78] H. DIGABEL et C. LANTUÉJOUL. Iterative algorithms. Dans J. L. CHERMANT, éditeur, *Proc. 2nd European Symp. Quantitative Analysis of Microstructures in Material Science, Biology and Medicine*, pages 85–99, Stuttgart, 1978. Riederer Verlag. [151](#)
- [DL01] L. DEVROYE et G. LUGOSI. *Combinatorial Methods in Density Estimation*. Springer, 2001. [32](#), [34](#)
- [DT77] L.P. DEVROYE et T.J.WAGNER. The strong uniform consistency of nearest neighbor density estimates. *Annals of Statistics*, 5(3) :536–540, 1977. [32](#)

- [DZ01] M.C. DELFOUR et J.P. ZOLÉSIO. *Shape and geometries*. Advances in Design and Control, SIAM, 2001. [19](#), [21](#)
- [EC87] C. EPSTEIN et M. CAGE. *The curve shortening flow*. Wave Motion : Theory, Modeling, and Computation, Springer-Verlag, 1987. [13](#)
- [Far99] G. FARNEBÄCK. Spatial domain methods for orientation and velocity estimation. Lic. Thesis LiU-Tek-Lic-1999 :13, Dept. EE, Linköping University, SE-581 83 Linköping, Sweden, mars 1999. Thèse No. 755, ISBN 91-7219-441-3. [101](#)
- [FLJ00] M.A.T. FIGUEIREDO, J.M.N. LEITAO, et A.K. JAIN. Unsupervised contour representation and estimation using b-splines and a minimum description length criterion. *IEEE Transactions on Image Processing*, 9(6) :1075–1087, juin 2000. [16](#), [17](#)
- [FM81] K. S. FU et J. K. MUI. A survey on image segmentation. *Pattern Recognition*, 13(1) :3–16, 1981. [2](#)
- [Fuk90] K. FUKUNAGA. *Introduction to statistical pattern recognition (2nd ed.)*. Academic Press Professional, Inc., New York, NY, USA, 1990. [61](#)
- [FZ04] D. FREEDMAN et T. ZHANG. Active contours for tracking distributions. *IEEE Transactions on Image Processing*, 13(4) :518–526, 2004. [17](#)
- [FZ05] D. FREEDMAN et T. ZHANG. Interactive graph cut based segmentation with shape priors. Dans *Computer Vision and Pattern Recognition (1)*, pages 755–762, 2005. [17](#)
- [GB02] M. GASTAUD et M. BARLAUD. Video segmentation using active contours on a group of pictures. Dans *IEEE International Conference on Image Processing*, volume 2, pages 81–84, Rochester, N.Y., septembre 2002. [100](#)
- [GBA04] M. GASTAUD, M. BARLAUD, et G. AUBERT. Combining shape prior and statistical features for active contour segmentation. *IEEE TCSVT special session on Audio and Video Analysis for Interactive Multimedia Services*, 14(5) :726–734, mai 2004. [87](#), [89](#)
- [GD05] R. GRAY et L. DAVISSON. *An introduction to statistical signal processing*. Cambridge University Press, 2005. [54](#)
- [GG84] S. GEMAN et D. GEMAN. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6) :721–741, novembre 1984. [47](#)
- [GLMI05] M.N. GORIA, N.N. LEONENKO, V.V. MERGEL, et P.L. Novi INVERARDI. A new class of random vector entropy estimators and its applications in testing statistical hypotheses. *Journal of Nonparametric Statistics*, 17(3) :277–297, 2005. [40](#), [119](#)
- [GML85] S. GEMAN et D.E. MC LURE. Bayesian image analysis : an application to single photon emission tomography. Dans *Proceedings of the Statistical Computing Section of the American Statistical Association*, 1985. [157](#)

- [GPS89] D. GREIG, B. PORTEOUS, et A. SEHEULT. Exact maximum a posteriori estimation for binary images. *Journal of Royal Statistical Society, Série B*, 51 :271–279, 1989. [17](#)
- [GR96] O. GERMAIN et Ph. RÉFRÉGIER. Optimal snake-based segmentation of a random luminance target on a spatially disjoint background. *Optics Letters*, 21(22) :1845–1848, novembre 1996. [16](#), [47](#)
- [GR04] F. GALLAND et Ph. RÉFRÉGIER. Information theory-based snake adapted to multi-region objects with different noise models. *Optic Letters*, 29(14) :1611–1614, 2004. [16](#), [17](#)
- [Gra90] R. M. GRAY. *Entropy and Information Theory*. Springer-Verlag, New York, New York, 1990. [39](#)
- [GRD04] F. GOUDAIL, Ph. RÉFRÉGIER, et G. DELYON. Bhattacharyya distance as a contrast parameter for statistical processing of noisy optical images. *Journal of the Optical Society of America A*, 21(7) :1231–1240, juillet 2004. [80](#)
- [Har75] J. A. HARDIGAN. *Clustering Algorithms*. John Wiley & Sons, New York, 1975. [5](#)
- [HJBD⁺06] A. HERBULOT, S. JEHAN-BESSON, S. DUFFNER, M. BARLAUD, et G. AUBERT. Segmentation of vectorial image features using shape gradients and information measures. *Journal of Mathematical Imaging and Vision*, 25(3) :365–386, octobre 2006. [100](#)
- [HP77] S.L. HOROWITZ et T. PAVLIDIS. Picture segmentation by a directed split and merge procedure. Dans *Computer Methods in Images Analysis*, pages 101–111, 1977. [5](#)
- [HR03] M. HINTERMÜLLER et W. RING. A second order shape optimization approach for image segmentation. *SIAM Journal on Applied Mathematics*, 64(2) :442–467, 2003. [21](#)
- [HS81] B. K. P. HORN et B. G. SCHUNCK. Determining optical flow. *Artificial Intelligence*, 17 :185–203, 1981. [100](#)
- [HS92] R.M. HARALICK et L.G. SHAPIRO. *Computer and Robot Vision*, volume 1. Addison-Wesley, 1992. [3](#), [4](#)
- [Hub81] P.J. HUBER. *Robust Statistics*. John Wiley and Sons, 1981. [118](#)
- [IAB⁺96] M. IRANI, P. ANANDAN, J. BERGEN, R. KUMAR, et S. HSU. Efficient representations of video sequences and their applications. *Signal Processing : Image Communication*, 8(4) :327–351, mai 1996. [100](#)
- [JB03] S. JEHAN-BESSON. *Modèles de contours actifs basés régions pour la segmentation d’images et de vidéos*. Thèse de doctorat, Université de Nice-Sophia Antipolis, 2003. [26](#)
- [JbBA02] S. JEHAN-BESSON, M. BARLAUD, et G. AUBERT. A 3-step algorithm using region-based active contours for video objects detection. *EURASIP journal on applied signal processing*, 2002(6) :572–581, 2002. [100](#), [155](#)

- [JBBA03] S. JEHAN-BESSON, M. BARLAUD, et 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. [17](#), [18](#), [21](#), [32](#)
- [JBU01] M. JACOB, T. BLU, et M. UNSER. A unifying approach and interface for spline-based snakes. Dans M. SONKA et K.M. HANSON, éditeurs, *Progress in Biomedical Optics and Imaging*, vol. 2, no. 27, volume 4322 de *Proceedings of the SPIE International Symposium on Medical Imaging : Image Processing*, pages 340–347, San Diego CA, USA, février 2001. Part I. [28](#)
- [Jef46] H. JEFFREYS. An invariant form for the prior probability in estimation problems). *Journal of Royal Statistical Society, Series A*, 186 :453–461, 1946. [72](#)
- [KDA99] P. KORNPORST, R. DERICHE, et G. AUBERT. Image sequence analysis via partial differential equations. *Journal of Mathematical Imaging and Vision*, 11 :5–26, 1999. [100](#)
- [KFY⁺05] J. KIM, J. W. FISHER, A. J. YEZZI, M. ÇETIN, et A. S. WILLSKY. A nonparametric statistical method for image segmentation using information theory and curve evolution. *IEEE Transactions on Image Processing*, 14(10) :1486–1502, 2005. [18](#), [47](#), [63](#)
- [KKO⁺95] S. KICHENASSAMY, A. KUMAR, P. OLVER, A. TANNENBAUM, et A. YEZZI. Gradient flows and geometric active contour models. Dans *International Conference on Computer Vision*, pages 810–815, Boston, USA, 1995. [15](#)
- [KKO⁺96] S. KICHENASSAMY, A. KUMAR, P. OLVER, A. TANNENBAUM, et A. YEZZI. Conformal curvature flows : from phase transitions to active vision. *Archive for Rational Mechanics and Analysis*, 134 :275–301, 1996. [15](#)
- [KL51] S. KULLBACK et R. A. LEIBLER. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1) :79–86, mars 1951. [42](#)
- [KL87] L. KOZACHENKO et N. LEONENKO. On statistical estimation of entropy of random vector. *Problems of Information Transmission*, 23 :95–101, 1987. [40](#)
- [KWT88] M. KASS, A. WITKIN, et D. TERZOPOULOS. Snakes : Active contour models. *International Journal of Computer Vision*, 1 :321–332, 1988. [x](#), [6](#), [13](#), [28](#)
- [LGF00] M. LEVENTON, E. GRIMSON, et O. FAUGERAS. Statistical shape influence in geodesic active contour. Dans *Computer Vision and Pattern Recognition*, pages 1316–1323, Hilton Head Island, South Carolina, 2000. [47](#), [146](#)
- [LK81] B.D. LUCAS et T. KANADE. An iterative image registration technique with an application to stereo vision. Dans *International Joint Conference on Artificial Intelligence*, pages 674–679, 1981. [100](#)
- [LRDF04] C. LENGLET, M. ROUSSON, R. DERICHE, et O. FAUGERAS. Statistics on multivariate normal distributions : A geometric approach and its applica-

- tion to diffusion tensor mri. Rapport technique 5242, INRIA, juin 2004. [47](#)
- [McQ67] J. MCQUEEN. Some methods for classification and analysis of multivariate observations. Dans *Proc. of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pages 281–297, 1967. [5](#)
- [MCV⁺97] F. MAES, A. COLLIGNON, D. VANDERMEULEN, G. MARCHAL, et P. SUE-TENS. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2) :187–198, 1997. [63](#)
- [MDM99] J. MONTAGNAT, H. DELINGETTE, et G. MALANDAIN. Cylindrical Echocardiographic Image Segmentation Based on 3D Deformable Models. Dans *International Conference on Medical Image Computing and Computer Assisted Intervention*, LNCS, pages 168–175, Cambridge, UK, septembre 1999. Springer-Verlag. [17](#)
- [MEFM02] A. MITICHE, R. EL-FEGHALI, et A.R. MANSOURI. Tracking moving objects as spatio-temporal boundary detection. Dans *Southwest Symposium on Image Analysis and Interpretation*, pages 106–110, 2002. [101](#)
- [Mey88] F. MEYER. *Skeletons in digital spaces*, chapitre Image Analysis and Mathematical Morphology. London : Academic Press, 1988. [151](#)
- [Min05] Thomas MINKA. Divergence measures and message passing. Rapport technique, Microsoft Research, 2005. [72](#)
- [MK99] A.R. MANSOURI et J. KONRAD. Motion segmentation with level sets. Dans *International Conference on Image Processing*, Kobe, Japon, 1999. [101](#)
- [MP98a] E. MÉMIN et P. PÉREZ. Dense estimation and object-based segmentation of the optical flow with robust techniques. *IEEE Transactions on Image Processing*, 7(5) :703–718, mai 1998. [110](#)
- [MP98b] E. MÉMIN et P. PÉREZ. Joint estimation-segmentation of optic flow. Dans *European Conference on Computer Vision1*, volume II, pages 563–577, Freiburg, Germany, juin 1998. [101](#)
- [MP04] A. MITTAL et N. PARAGIOS. Motion-based background subtraction using adaptive kernel density estimation. Dans *Computer Vision and Pattern Recognition*, volume 2, pages 302–309, 2004. [119](#)
- [MRGG06] P. MARTIN, P. RÉFRÉGIER, F. GALLAND, et F. GUÉRAULT. Nonparametric statistical snake based on the minimum stochastic complexity. *IEEE Transactions on Image Processing*, 15(9) :2762–2770, 2006. [16](#), [17](#), [47](#)
- [MRT07] O. MICHAILOVICH, Y. RATHI, et A. TANNENBAUM. Image segmentation using active contours driven by information-based criteria. *IEEE Transactions on Image Processing*, à paraître, 2007. [80](#)
- [MS89] D. MUMFORD et J. SHAH. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42 :577–684, 1989. [17](#), [47](#)

- [MSMM90] S. MENET, P. SAINT MARC, et G. MEDIONI. B-snakes : Implementation and application to stereo. Dans *Image Understanding Workshop*, pages 720–726, 1990. [28](#)
- [MSV95] R. MALLADI, J.A. SETHIAN, et B.C VEMURI. Shape modeling with front propagation : a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17 :158–175, février 1995. [15](#), [28](#)
- [MSV96] R. MALLADI, J.A. SETHIAN, et B.C. VEMURI. A fast level set based algorithm for topology independant shape modeling. *Journal of Mathematical Imaging and Vision, Special issue on Topology and Geometry*, 6 :269 – 289, 1996. A. Rosenfeld and Y.Kong, eds. [15](#)
- [NBN07] F. NIELSEN, J.-D. BOISSONNAT, et R. NOCK. On bregman voronoi diagrams. Dans *ACM-SIAM Symposium on Discrete Algorithms*, pages 746–755, 2007. [145](#)
- [OB95] J.-M. ODOBEZ et P. BOUTHEMY. Robust multiresolution estimation of parametric motion models. *Journal of Visual Communication and Image Representation*, 6 :348–365, 1995. [100](#), [118](#), [120](#), [155](#)
- [OS88] S. OSHER et J.A. SETHIAN. Fronts propagating with curvature-dependent speed : Algorithms based on hamilton-jacobi formulation. *Journal of Computational Physics*, 79 :12–49, 1988. [15](#), [26](#)
- [Par62] E. PARZEN. On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 33 :1065–1076, 1962. [32](#)
- [PBBU05] F. PRECIOSO, M. BARLAUD, T. BLU, et M. UNSER. Robust real-time segmentation of images and videos using a smooth-spline snake-based algorithm. *IEEE Transactions on Image Processing*, 14(7) :910–924, juillet 2005. [17](#), [28](#)
- [PBMR00] C. PAPIN, P. BOUTHEMY, E. MÉMIN, et G. ROCHARD. Tracking and characterization of highly deformable cloud structures. Dans *European Conference on Computer Vision, ECCV 2000, LNCS 1843*, volume 2, pages 428–442, Dublin, Irlande, juin 2000. [100](#)
- [PD99] N. PARAGIOS et R. DERICHE. Geodesic active regions for motion estimation and tracking. Dans *International Conference on Computer Vision*, Corfu Greece, 1999. [100](#)
- [PD00] N. PARAGIOS et R. DERICHE. Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22 :266–280, mars 2000. [17](#), [99](#)
- [PD02] N. PARAGIOS et R. DERICHE. Geodesic active regions : a new paradigm to deal with frame partition problems in computer vision. *Journal of Visual Communication and Image Representation, Special Issue on Partial Differential Equations in Image Processing, Computer Vision and Computer Graphics*, 13(1/2) :249–268, mars/juin 2002. [47](#)

- [PD05] N. PARAGIOS et R. DERICHE. Geodesic active regions and level set methods for motion estimation and tracking. *Computer Vision and Image Understanding*, 97(3) :259–282, mars 2005. [47](#)
- [PF00] E. J. PAUWELS et G. FREDERIX. Image segmentation by nonparametric clustering based on the kolmogorov-smirnov distance. Dans *Proceedings of the 6th European Conference on Computer Vision - Part II*, pages 85–99, London, UK, 2000. Springer-Verlag. [80](#)
- [RAB06] C. C. REYES-ALDASORO et A. BHALERAO. The bhattacharyya space for feature selection and its application to texture segmentation. *Pattern Recognition*, 39(5) :812–826, 2006. [80](#)
- [RBD03] M. ROUSSON, T.S BROX, et R. DERICHE. Active unsupervised texture segmentation on a diffusion based feature space. Dans *Conference on Computer Vision and Pattern Recognition*, volume 02, pages 699–704, Los Alamitos, CA, USA, juin 2003. IEEE Computer Society. [145](#)
- [RD02] M. ROUSSON et R. DERICHE. A variational framework for active and adaptive segmentation of vector valued images. Dans *Proc. IEEE Workshop on Motion and Video Computing*, pages 56–62, Orlando, Florida, décembre 2002. [47](#)
- [RD04] F. RANCHIN et F. DIBOS. Moving objects segmentation using optical flow estimation. Dans *Mathematics, Image and Analysis*, Paris, France, 2004. [100](#)
- [RD05] F. RANCHIN et F. DIBOS. Variational level set methods : from continuous to discrete setting, applications in video segmentation and tracking. Dans *IEEE International Conference on Image Processing*, volume 1, pages 273–276, 2005. [100](#)
- [Réf05] P. RÉFRÉGIER. Polarization degree of optical waves with non-Gaussian probability density functions : Kullback relative entropy-based approach. *Optics Letters*, 30 :1090–1092, mai 2005. [72](#)
- [RG06] P. RÉFRÉGIER et F. GOUDAIL. Kullback relative entropy and characterization of partially polarized optical waves. *Journal of the Optical Society of America A*, 23 :671–678, mars 2006. [72](#)
- [RJZ06] M. ROCHERY, I. H. JERMYN, et J. ZERUBIA. Higher order active contours. *International Journal of Computer Vision*, 69 :27–42, 2006. [47](#)
- [RKB04] C. ROTHER, V. KOLMOGOROV, et A. BLAKE. "grabcut" : interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics*, 23(3) :309–314, 2004. [17](#)
- [RM01] J.B.T.M. ROERDINK et A. MEIJSTER. The watershed transform : Definitions, algorithms and parallelization strategies. *Fundamenta Informaticae*, 41 :187–228, 2001. [151](#)
- [Ron94] R. RONFARD. Region-based strategies for active contour models. *International Journal of Computer Vision*, 13(2) :229–251, 1994. [16](#)

- [Ros56] M. ROSENBLATT. Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, 27 :832–837, 1956. [32](#)
- [RP02] M. ROUSSON et N. PARAGIOS. Shape priors for level set representations. Dans *European Conference in Computer Vision*, pages 78–92, Copenhagen, Danemark, 2002. [146](#)
- [RR01] O. RUCH et Ph. RÉFRÉGIER. Minimal-complexity segmentation with a polygonal snake adapted to different optical noise models. *Optical Letters*, 41 :977–979, 2001. [16](#), [17](#)
- [RTG00] Y. RUBNER, C. TOMASI, et L. J. GUIBAS. The earth mover’s distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2) :99–121, 2000. [80](#)
- [San57] I. N. SANOV. On the probabilities of large deviations of random variables. *Selected Translations in Mathematical Statistics and Probability*, 1 :214–244, 1957. [43](#)
- [SBFAZ00] C. SAMSON, L. BLANC-FÉRAUD, G. AUBERT, et J. ZERUBIA. A level set model for image classification. *International Journal of Computer Vision*, 40(3) :187–197, 2000. [47](#)
- [Sch94] C. SCHNÖRR. Segmentation of visual motion by minimizing convex non-quadratic functionals. Dans *Proc. 12th International Conference on Pattern Recognition*, volume A, pages 661–663, Jérusalem, octobre 1994. IEEE Computer Society Press, Los Alamitos. [110](#)
- [Sco92] D.W. SCOTT. *Multivariate Density Estimation : Theory, Practice, and Visualization*. Wiley, 1992. [32](#), [34](#), [119](#), [145](#)
- [Ser06] J. SERRA. A lattice approach to image segmentation. *Journal of Mathematical Imaging and Vision*, 24(1) :83–130, 2006. [2](#)
- [Set96] J.A. SETHIAN. *Level Set Methods*. Cambridge Univ. Press, 1996. [26](#), [28](#)
- [SG00] P. SALEMBIER et L. GARRIDO. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Trans. on Image Processing*, 9(4) :avril, 2000. [153](#)
- [Sha48] C.E. SHANNON. A mathematical theory of communication. *Bell System Technical Journal*, 27 :379–423 and 623–656, juillet et octobre 1948. [31](#), [39](#)
- [Sil86] B.W. SILVERMAN. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, Londres, 1986. [32](#), [34](#), [119](#), [126](#)
- [SK94] W. SKARBEEK et A. KOSCHAN. Colour image segmentation — a survey. Rapport technique, Institute for Technical Informatics, Technical University of Berlin, octobre 1994. [2](#)
- [SM99] P. SALEMBIER et F. MARQUÈS. Region-based representations of image and video : segmentation tools for multimedia services. *IEEE Transactions on Circuits and Systems for Video Technology*, 9(8) :1–20, décembre 1999. [100](#)

- [SSWC88] P. K. SAHOO, S. SOLTANI, A. K.C. WONG, et Y. C. CHEN. A survey of thresholding techniques. *Computer Vision, Graphics, and Image Processing*, 41(2) :233–260, 1988. [3](#)
- [ST01] E. SIFAKIS et G. TZIRITAS. Moving object localisation using a multi-label fast marching algorithm. *Signal Processing : Image Communication*, 16 :963–976, 2001. [100](#)
- [SY02] S. SOATTO et A. J. YEZZI. Deformation : Deforming motion, shape average and the joint registration and segmentation of images. Dans *European Conference on Computer Vision*, pages 32–57, 2002. [101](#)
- [SZ92] J. SOKOLOWSKI et J.P. ZOLESIO. *Introduction to shape optimization*, volume 16 de *Springer series in computational mathematics*. Springer-Verlag, 1992. [19](#), [21](#)
- [TJW01] A. TSAI, A. J. Yezzi JR., et A. S. WILLSKY. Curve evolution implementation of the mumford-shah functional for image segmentation, denoising, interpolation, and magnification. *IEEE Transactions on Image Processing*, 10(8) :1169–1186, 2001. [47](#)
- [TRRK98] J.Y. THAM, S. RANGANATH, M. RANGANATH, et A.A. KASSIM. A novel unrestricted center-biased diamond search algorithm for block motion estimation. *IEEE Transactions on Circuits and Systems for Video Technology*, 8(4) :369–377, août 1998. [127](#)
- [TYW⁺03] A. TSAI, A. YEZZI, W. WELLS, C. TEMPANY, D. TUCKER, A. FAN, W.E. GRIMSON, et A. WILLSKY. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 22(2) :137–154, 2003. [146](#)
- [Tür94] B.A. TÜRLACH. Fast implementation of density-weighted average derivative estimation. *Computationally Intensive Statistical Methods*, 26 :28–33, 1994. [33](#)
- [Uns99] M. UNSER. Splines : A perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, 16(6) :22–38, novembre 1999. [28](#)
- [Vas76] O. VASICEK. A test for normality based on sample entropy. *Journal of Royal Statistical Society, Series B*, 38(1) :54–59, 1976. [40](#)
- [VC02] L. A. VESE et T. F. CHAN. A multiphase level set framework for image segmentation using the mumford and shah model. *International Journal of Computer Vision*, 50(3) :271–293, 2002. [47](#)
- [VML06] C. VAZQUEZ, A. MITICHE, et R. LAGANIERE. Joint multiregion segmentation and parametric estimation of image motion by basis function representation and level set evolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(5) :782–793, 2006. [101](#)
- [VS91] L. VINCENT et P. SOILLE. Watersheds in digital spaces : An efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6) :583–598, 1991. [151](#)

- [VW97] P. VIOLA et W. M. WELLS. Alignment by Maximization of Mutual Information. *International Journal of Computer Vision*, 24(2) :137–154, 1997. [63](#)
- [Wei98] J. WEICKERT. On discontinuity-preserving optic flow. Dans *Proc. Computer Vision and Mobile Robotics Workshop, Santorini*, pages 115–122, septembre 1998. [110](#)
- [WK93] S. F. WU et J. KITTLER. A gradient-based method for general motion estimation and segmentation. *Journal of Visual Communication and Image Representation*, 4(1) :25–38, mars 1993. [100](#), [117](#), [131](#)
- [WS01] J. WEICKERT et C. SCHNÖRR. Variational optic flow computation with a spatio-temporal smoothness constraint. *Journal of Mathematical Imaging and Vision*, 14 :245–255, 2001. [100](#)
- [WTP05] E. WOLSZTYNSKI, E. THIERRY, et L. PRONZATO. Minimum entropy estimation in semi parametric models. *IEEE Signal Processing*, 85 :937–949, 2005. [40](#)
- [WV04] Z. WANG et Baba C. VEMURI. An affine invariant tensor dissimilarity measure and its applications to tensor-valued image segmentation. Dans *Computer Vision and Pattern Recognition*, 2004. [47](#)
- [WVA+96] W. WELLS, P. VIOLA, H. ATSUMI, S. NAKAJIMA, et R. KIKINIS. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1) :35–51, 1996. [63](#)
- [XBA03] N. XU, R. BANSAL, et N. AHUJA. Object segmentation using graph cuts based active contours. Dans *Computer Vision and Pattern Recognition*, pages 46–53, juin 2003. [17](#)
- [YTW99] A. YEZZI, A. TSAI, et A. WILLSKY. A statistical approach to snakes for bimodal and trimodal imagery. Dans *International Conference on Image Processing*, pages 898–903, Kobe Japon, 1999. [17](#), [32](#), [47](#)
- [ZCMO96] H.-K. ZHAO, T. CHAN, B. MERRIMAN, et S. OSHER. A variational level set approach to multiphase motion. *Journal of Computational Physics*, 127(1) :179–195, 1996. [47](#)
- [Zha01] D. ZHANG. Segmentation of moving objects in image sequence : A review. *Circuits, Systems, and Signal Processing*, 20(2) :143–183, 2001. [99](#)
- [ZK04] R. ZABIH et V. KOLMOGOROV. Spatially coherent clustering using graph cuts. Dans *Computer Vision and Pattern Recognition*, pages 437–444, juin 2004. [17](#)
- [ZY96] S. ZHU et A. YUILLE. Region competition : unifying snakes, region growing, and bayes/MDL for multiband image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18 :884–900, septembre 1996. [16](#), [47](#), [51](#)

Publications

REVUES INTERNATIONALES

IJCV'08, S. Boltz, A. Herbulot, E. Debreuve, M. Barlaud, G. Aubert. "Motion and appearance nonparametric joint entropy for video segmentation", *International Journal of Computer Vision*, accepté pour publication, publié en ligne le 19 janvier 2008.

JMIV'06, A. Herbulot, S. Jehan-Besson, S. Duffner, M. Barlaud, G. Aubert. "Segmentation of vectorial image features using shape gradients and information measures", *Journal of Mathematical Imaging and Vision*, Volume 25, Issue 3, pp 365-386, octobre 2006.

CHAPITRES DE LIVRE

SPRINGER'05, S. Jehan-Besson, A. Herbulot, M. Barlaud, G. Aubert. "Shape Gradient for Image and Video Segmentation", *Mathematical Models in Computer Vision : The Handbook*, éditeurs Nikos Paragios, Yunmei Chen et Olivier Faugeras, 2005.

CONFÉRENCES INTERNATIONALES

SSVM'07, A. Herbulot, S. Boltz, E. Debreuve, M. Barlaud, G. Aubert. "Space-time segmentation based on a joint entropy with estimation of nonparametric distributions", *International Conference on Scale Space Methods and Variational Methods in Computer Vision*, LNCS, Vol. 4485, pp 721-732, Ischia, Italie, mai 2007.

ICIP'06, A. Herbulot, S. Boltz, E. Debreuve, M. Barlaud. "Robust motion-based segmentation in video sequences using entropy estimator", *International Conference on Image Processing*, pp 1853-1856, Atlanta, USA, octobre 2006.

SMVP'06, S. Boltz, A. Herbulot, E. Debreuve, M. Barlaud. "Entropy-based space-time segmentation in video sequences", *Workshop on Statistical Methods in Multi-Image and Video Processing, in conjunction with ECCV 2006*, Graz, Autriche, mai 2006. **Best student paper.**

ICIP'05, A. Gouze, C. De Roover, A. Herbulot, E. Debreuve, M. Barlaud, B. Macq. "Watershed-driven Active Contours for Moving Object Segmentation", *International Confe-*

rence on Image Processing, vol II, pp 818-821, Genova, Italie, septembre 2005.

EUSIPCO'05, C. De Roover, A. Herbulot, A. Gouze, E. Debreuve, M. Barlaud, B. Macq. "Multimodal Segmentation Combining Active Contours and Watersheds", *European Signal Processing Conference*, Antalya, Turquie, septembre 2005.

ICIP'04, A. Herbulot, S. Jehan-Besson, M. Barlaud, G. Aubert. "Shape gradient for multimodal image segmentation using mutual information", *International Conference on Image Processing*, pp 2729-2732, Singapour, octobre 2004.

ICASSP'04, A. Herbulot, S. Jehan-Besson, M. Barlaud, G. Aubert. "Shape gradient for image segmentation using information theory", *International Conference on Acoustics, Speech, and Signal Processing*, vol III, pp 21-24, Montreal, Canada, mai 2004.

WIAMIS'04, A. Herbulot, S. Jehan-Besson, M. Barlaud, G. Aubert. "Shape gradient for multimodal image segmentation using joint intensity distributions", *Workshop on Image Analysis for Multimedia Interactive Services*, Lisbonne, Portugal, avril 2004.

CONFÉRENCES NATIONALES

GRETSI'05, S. Jehan-Besson, S. Duffner, A. Herbulot, M. Barlaud, G. Aubert. "Utilisation des gradients de forme et des contours actifs basés régions pour la segmentation des vecteurs mouvement", *GRETSI conférence sur le traitement du signal et des images*, Louvain la Neuve, Belgique, septembre 2005.

AUTRES CONFÉRENCES

MIA'04, S. Jehan-Besson, A. Herbulot, M. Barlaud, G. Aubert. "Shape gradients for image and video segmentation", *Mathematics and Image Analysis*, Paris, septembre 2004.

RAPPORTS DE RECHERCHE

Rapport de DEA, A. Herbulot, "Segmentation d'images et de vidéo en utilisant des histogrammes", juin 2003.

Rapport de recherche interne I3S, RR-2004-43-FR, A. Herbulot, S. Jehan-Besson, M. Barlaud, G. Aubert. "Information theory for image segmentation using shape gradient", décembre 2004.

RÉSUMÉ

La segmentation d'images et de séquences vidéo consiste à séparer les objets d'intérêt du reste de l'image. L'utilisation de contours actifs en utilisant une approche variationnelle repose sur la définition d'un critère de segmentation. Ce critère est ensuite dérivé à l'aide des gradients de forme, afin de parvenir à l'équation d'évolution du contour actif. Bien souvent ce critère dépend des caractéristiques de l'image et fait une hypothèse implicite sur les distributions de ces caractéristiques. Par exemple, considérer une fonction de la moyenne de l'intensité d'une région revient à faire une hypothèse gaussienne sur la distribution de cette intensité.

Dans cette thèse, nous proposons de nous affranchir de ces hypothèses qui ne sont pas toujours respectées et de considérer les distributions les plus "réelles" possible en utilisant une estimation non-paramétrique de ces distributions. Nous présentons des critères issus de la théorie de l'information, comme l'entropie, afin de segmenter des zones de faible variabilité dans les images. Afin de prendre en compte plusieurs canaux comme les canaux couleur, l'entropie jointe et l'information mutuelle sont aussi utilisées. Lorsqu'une information *a priori* est connue, la divergence de Kullback-Leibler permet d'introduire une notion de distance à une segmentation de référence en cherchant à minimiser une "distance" entre distributions. Enfin, l'entropie jointe est utilisée afin de segmenter des objets en mouvement dans des séquences vidéo, que cela soit en ayant au préalable calculé un flot optique, ou en estimant de façon conjointe le mouvement avec la segmentation.

Mots clés : Segmentation, contours actifs, méthode variationnelle, équation aux dérivées partielles, gradient de forme, théorie de l'information, entropie, entropie jointe, information mutuelle, divergence de Kullback-Leibler, estimation de mouvement.

ABSTRACT

Image and video segmentation consists in the partitioning of an image into objects of interest and background. When using active contours in an variational framework, the difficulty is to define an appropriate segmentation criterion. This criterion is then differentiated using shape gradients, in order to obtain the evolution equation of the active contour. Often this criterion depends on image features and makes an assumption on the distribution of such features. For example, considering a function of the intensity mean as a criterion is equivalent to making a Gaussian assumption on the distribution of the intensity.

In this work, we propose to get rid of such assumptions by approximating actual distributions. We use a non-parametric kernel-based estimator. We propose different criteria coming from information theory, such as entropy, to segment zones with limited intensity variations. In order to take into account several channels like color channels, two alternatives are proposed : joint entropy and mutual information. When some *a priori* is available, the Kullback-Leibler divergence is used to minimize a distance between a reference distribution and the distribution of the current region. To segment moving objects in video sequences, the joint entropy is used. A first approach consists in computing the optical flow and minimizing the joint entropy of its components. A second approach consists in jointly estimating the motion and segmenting moving objects by minimizing the joint entropy of a residual and the image intensity.

Keywords : Segmentation, active contours, variational framework, partial differential equations, shape gradient, information theory, entropy, joint entropy, mutual information, Kullback-Leibler divergence, motion estimation.