Méthodes par ensembles de niveaux et modes conditionnels itérés pour la segmentation vidéo

Florent Ranchin

▶ **To cite this version:**

Florent Ranchin. Méthodes par ensembles de niveaux et modes conditionnels itérés pour la segmentation vidéo. Interface homme-machine [cs.HC]. Université Paris Dauphine - Paris IX, 2004. Français. NNT : . tel-00008038v2

HAL Id: tel-00008038

https://theses.hal.science/tel-00008038v2

Submitted on 4 Mar 2005

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 Paris Dauphine
UFR Mathématiques de la Décision

THÈSE

pour l'obtention du titre de

DOCTEUR EN SCIENCES
Spécialité Mathématiques Appliquées
(arrêté du 30 mars 1992)

présentée et soutenue par

Florent RANCHIN

Méthodes par ensembles de niveaux et modes conditionnels itérés pour la segmentation vidéo

Directeur de Recherches :
Mme Françoise DIBOS

JURY

Président :	**M. Jean-Michel MOREL** Professeur, Ecole Normale Supérieure de Cachan
Rapporteurs :	**M. Michel BARLAUD** Professeur, Université de Nice-Sophia Antipolis
	M. Frédéric CAO Chargé de Recherche INRIA habilité, IRISA.
Examinateurs :	**M. Laurent COHEN** Directeur de Recherche CNRS, Université Paris-Dauphine
	Mme Françoise DIBOS Maître de Conférences habilitée, Université Paris-Dauphine
	M. Richard LENGAGNE Chercheur, Ecole Polytechnique Fédérale de Lausanne
Invité :	**M. Jean-Michel MARIN** Maître de Conférences, Université Paris-Dauphine.

10 décembre 2004

Méthodes par ensembles de niveaux et modes conditionnels itérés pour la segmentation vidéo

Résumé : Cette thèse est consacrée à l'étude d'un problème de vision par ordinateur et de deux problèmes de vidéo surveillance.
Nous proposons une méthode de détection d'objets en mouvement dans une séquence vidéo basée sur une détermination préalable du mouvement apparent et sur un problème d'optimisation de forme.
Pour d'autres modèles de détection et de suivi d'objets en mouvement, nous proposons d'appliquer l'algorithme discret des modes conditionnels itérés réputé très rapide et qui permet de réduire le temps de calcul des algorithmes continus lorsqu'il leur est combiné.
En vidéo surveillance, on cherche d'une part à estimer la densité d'une foule et d'autre part à détecter des comportements anormaux dans l'environnement du métro parisien. Nous proposons une estimation de la densité d'une foule basée sur un calcul de courbure sur l'image. La détection de comportements anormaux s'effectue par une recherche des modes dans l'histogramme des directions du mouvement apparent.

Mots clés : analyse d'images et de séquences vidéo, méthodes variationnelles, méthodes par ensembles de niveaux, algorithmes de relaxation déterministe, flot optique.

Level Sets and Iterated Conditional Modes Methods for Video Segmentation

Abstract : This thesis is devoted to the study of a computer vision problem and of two video surveillance problems.
We propose a method for moving objects detection in a video sequence ; it is based on the apparent velocity predetermination and on a shape optimisation problem.
For other moving objects detection and tracking models, we propose to use the iterated conditional modes discrete algorithm which is renowned to be very fast and allows to decrease the computing time of the continuous algorithms when it is combined to them.
For video surveillance, we seek to estimate crowd density and to detect abnormal behaviour in parisian subway environment. We propose a crowd density estimation based on the computation of curvature on the image. Abnormal behaviour detection is performed by a research of the modes in the histogram of the apparent velocity directions.

Keywords : image and video analysis, variational methods, level sets methods, deterministic relaxation algorithms, optical flow.

A mes parents,

Remerciements

Jean-Michel Morel a accepté de présider ce jury. Je le remercie de l'honneur qui m'est ainsi fait.

Michel Barlaud et Frédéric Cao ont accepté d'être rapporteurs de cette thèse. Je les remercie pour tous leurs commentaires, remarques et suggestions qui furent fructueux et bénéfiques.

Je tiens à remercier Françoise Dibos pour m'avoir guidé dans mes recherches pendant ces trois ans. Grâce à elle, j'ai énormément appris. Ses qualités scientifiques, son enthousiasme et son soutien ont été une source constante de motivation.

Laurent Cohen et Richard Lengagne ont accepté de faire partie de ce jury. Je les en remercie vivement.

Je tiens à remercier Jean-Michel Marin de faire partie de ce jury en tant qu'invité.

Une partie des travaux présentés dans cette thèse a été effectuée dans le cadre d'un contrat CEREMADE-RATP-VISIOWAVE. Toute ma reconnaissance va à l'ensemble des personnels de la RATP et de la société VISIOWAVE que j'ai pu cotoyer, tout particulièrement Richard Lengagne et Philippe Bruneaut avec qui il aura été particulièrement agréable de travailler.

Je tiens aussi à remercier tout particulièrement Agnès Desolneux pour m'avoir invité à participer aux groupes de travail du laboratoire MAP5 de l'université Paris V. Mes remerciements s'adressent aussi à Jacques Froment qui m'a invité à présenter mes travaux au MIG 2004 organisé par le laboratoire LMAM de l'université Bretagne Sud. Ces remerciements s'adressent également à l'ensemble des chercheurs en traitement d'image rencontrés lors de ces occasions : Isabelle Abraham, Romain Abraham, Sylvain Durand, Yann Gousseau, Georges Koepfler, François Malgouyres, Simon Masnou, Lionel Moisan, Mila Nikolova, Frédéric Richard, Jian-Feng Yao et tous ceux que j'ai oubliés (et qui je l'espère se reconnaitront).

Je remercie également Laurent Cohen pour m'avoir invité à participer au MIA 2004 et m'avoir ainsi permis de valoriser mes recherches lors d'une conférence internationale.

Cette thèse a été effectuée au CEREMADE dirigé par Maria Esteban. Que l'ensemble des membres du CEREMADE, secrétaires, ingénieurs et chercheurs, soient ici remerciés pour leur accueil, tout particulièrement Josette Lévy et Magali Marc-Dibildos pour leur disponibilité et leur bonne humeur, Michel Vanbreugel, Gilles Barès, Fabrice Rossi et Antonin Chambolle pour leur aide précieuse en informatique... Ce fut un plaisir de travailler dans un tel cadre.

Je tiens à remercier vivement Christian Lopez, mon tuteur de monitorat, qui a guidé mes premiers pas dans l'enseignement. J'ai été ravi de pouvoir enseigner avec lui. Je le remercie pour son enthousiasme constant et son soutien tant pour mes enseignements que pour mes recherches. Toute ma reconnaissance va aussi à Claudine Dhuin et Filipe Da Rocha avec qui j'ai également enseigné dans le cadre de mon monitorat. Je remercie plus largement l'ensemble des personnels de l'UFR MD et plus particulièrement Denis Pasquignon pour toutes les discussions que nous avons eues autour d'un café...

Toute mon amitié va à mes amis et collègues thésards de Dauphine, Cachan et d'ailleurs : Jean-Pascal, Hycham, Feth, Stéphane, Nabile, Valérie, Omar, Ridha, Claire, Mathieu, Nora, Sylvain, Lassana, Olivia, Niousha, Amina ; Gwendoline, David, Julie, Jérôme, Bénédicte, Pablo, Amandine, Béranger, Frédéric, Sébastien, Thomas et tous ceux que j'ai oubliés (et qui je l'espère se reconnaitront).

Enfin, je remercie mes parents qui m'ont supporté pendant ces trois ans !...

Table des matières

1 Introduction **1**

1.1 Du pixel à la forme 1
1.2 Problématiques des contours actifs 2
1.3 Problématiques liées à la modélisation du mouvement 4
1.4 Solutions proposées 7
1.4.1 Segmentation et tracking d'objets en mouvement 7
1.4.1.1 Utilisation du flot optique 7
1.4.1.2 Optimisation de forme et calcul des variations . 8
1.4.1.3 Application de l'ICM à une classe de problèmes d'ensembles de niveaux 11
1.4.2 Problèmes de vidéo surveillance 11
1.5 Plan du mémoire 12

2 Méthodes de segmentation d'images fixes **15**

2.1 Présentation des contours actifs 15
2.1.1 Le Principe 15
2.1.2 Le premier modèle variationnel de Kass, Witkin et Terzopoulos 16
2.1.3 Le modèle variationnel de Caselles, Kimmel et Sapiro . . 17
2.1.4 Le modèle d'évolution de Caselles, Catté, Coll et Dibos . 18
2.1.4.1 Formulation scalaire 18
2.1.4.2 Dualité entre évolutions scalaire et géométriques, méthode d'Osher-Sethian 19
2.1.4.3 Formulation géométrique 20
2.2 Un modèle de contours actifs basés régions généralisant le modèle de Mumford et Shah 20
2.2.1 Formulation locale 21
2.2.2 Formulation scalaire 23
2.2.2.1 Réécriture de l'énergie 23
2.3 Résultats et comparaison des méthodes 25
2.4 Revue des implémentations 29
2.4.1 A partir du schéma d'Alvarez et Guichard 29
2.4.2 A partir du schéma de Malik et Perona 30

2.4.3 A partir d'opérateurs morphologiques 32
2.4.4 Implémentation des EDPs avec fonctions de Heaviside . . 33

3 Segmentation d'objets en mouvement à caméra fixe 35
3.1 Segmentation conjointe fond/objets en mouvement 35
3.2 Application de l'ICM et des contours actifs à la segmentation et au tracking d'objets en mouvement 40
3.2.1 Comparaison des images avec un fond donné 41
3.2.1.1 Formulation continue, contours actifs basés régions, EDPs 41
3.2.1.2 Formulation discrète, champs de Markov, ICM . 42
3.2.1.3 ICM vs EDP : temps de calcul et résultats . . . 45
3.2.1.4 Remarque sur le modèle d'Ising et la "longueur du chauffeur de taxi" 46
3.2.2 Suivi d'objets en mouvement : Une méthode de tracking sans modèle de mouvement *a priori* 47
3.2.2.1 Initialisation 52
3.2.2.2 Modèle continu 54
3.2.2.3 Modèle discret 55
3.2.2.4 Avantages et inconvénients de la méthode de tracking 2D 57
3.3 Estimation du flot optique et segmentation d'objets en mouvement par utilisation du flot optique 66
3.3.1 Le problème du flot optique 66
3.3.2 Remarque sur le flot optique avec multirésolution 68
3.3.3 Quelques modèles de segmentation par le flot optique . . 69
3.4 Un nouveau modèle de segmentation d'objets en mouvement par le flot optique 70
3.4.1 Le modèle variationnel 70
3.4.2 Analyse des différents termes de l'énergie et résultats expérimentaux 73

4 Second cas : la caméra est mobile 83
4.1 Calcul du mouvement de caméra 83
4.2 Soustraction du mouvement induit par le déplacement de la caméra et segmentation 85
4.2.1 Segmentation par comparaison entre deux images successives 85
4.2.2 Segmentation par comparaison entre l'image et le fond . . 86
4.3 Segmentation par utilisation du flot optique résiduel 91

5 Analyse vidéo en temps réel 105
5.1 Détermination des objets en mouvement par détection de changement 106
5.2 Méthodes locales de flot optique 108

5.2.1 La formulation de Lucas et Kanade 109
5.2.2 Introduction d'un filtrage 110
5.2.3 Utilisation des ondelettes (C. Bernard, [Ber]) 110
5.3 Comparaison avec une méthode différentielle 111
5.4 Expériences, critère d'anormalité 111
5.5 Estimation de la densité d'une foule 115

A Théorie de la dérivation de domaine **127**
A.1 Dérivées d'énergies classiques 130
A.1.1 Longueur, aire et géodésiques 130
A.1.2 Energies impliquant moyennes et variances 131
A.2 Dérivée de $\boldsymbol{J(\Gamma) = \int_\Gamma g(\nabla u \cdot \mathbf{n})\, d\mathcal{H}^1}$ 133

B Documentation des modules MegaWave utilisés **137**

Notations

- Les vecteurs et les points du plan sont notés en caractères gras.
- $\nabla = (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2})^T$ désigne l'opérateur gradient spatial (en deux dimensions) et $\nabla^\theta = (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial t})^T$ désigne l'opérateur gradient spatio-temporel.
- Une image est représentée par la fonction continue $I : D \rightarrow \mathbb{R}$. $I(\mathbf{x})$ représente donc le niveau de gris de l'image au point (ou pixel) $\mathbf{x}$. Lorsque l'on fera référence à des modèles discrets, l'image sera toujours notée I et correspond à la discrétisation de l'image sur une grille régulière.
- Une séquence d'images sera notée (I_t) où $I_t : D \rightarrow \mathbb{R}$ où t appartient à l'intervalle d'entiers $0, 1, ..., T$. La dérivée temporelle sera notée $\dfrac{\partial}{\partial t}$ ou ∂_t pour éviter les confusions.
- Dans les problèmes d'optimisation de forme, le domaine recherché sera noté Ω et sa frontière $\Gamma \triangleq \partial\Omega$. Son complémentaire $D \setminus \Omega$ dans D sera noté Ω^c et sa normale extérieure sera désignée par $\mathbf{n}$. La courbure de Γ au point $\mathbf{x}$ est notée $\kappa(\mathbf{x})$ ou plus simplement κ lorsqu'il n'y a pas d'ambiguïté. Afin de ne pas alourdir les notations, on notera les fonctionnelles $J(\Omega)$ ou $J(\Gamma)$ plutôt que $J(\Omega, \Gamma)$.
- La mesure de Hausdorff de dimension $N-1$ dans $\mathbb{R}^N$ est notée $\mathcal{H}^{N-1}$. Elle est souvent utilisée dans les problèmes d'optimisation de forme faisant intervenir un terme de contour dans le cas $N = 2$ ou d'hypersurface dans le cas $N > 2$.
- Bien souvent, on utilisera l'abréviation EDP pour dire Equations aux Dérivées Partielles et ICM pour Iterated Conditional Modes.
- Dans le texte, les références aux figures sont faites **en gras**.

Chapitre 1

Introduction

1.1 Du pixel à la forme

L'analyse d'une image ou d'une vidéo est une opération simple pour le cerveau humain. Pourtant, elle semble très difficile à modéliser mathématiquement et à imiter par des outils informatiques et automatiques. Pourquoi ? parce que comme cela a été expliqué par M. Merleau-Ponty dans la théorie de la forme, l'analyse visuelle des formes par le cerveau s'effectue à un niveau global et non local comme tendrait à le faire basiquement un système automatique.
En effet, lorsque l'on traite informatiquement une image, la première donnée à laquelle on a accès est le niveau de gris ou la couleur d'un pixel. La difficulté est donc de remonter de cette information très "bas niveau" vers une information plus riche. C'est à cette difficulté que l'on est directement confronté lorsque l'on a affaire à un problème de *segmentation*. Parmi les différents problèmes de segmentation, on peut considérer les suivants : segmentation de l'image en régions homogènes, segmentation d'objets en mouvement dans une vidéo ou encore suivi d'objets ou *tracking* à partir de la donnée des objets sur la première image de la vidéo. Ces différents problèmes trouvent leurs applications dans la compression vidéo, la réalité augmentée ou l'analyse de séquences d'imagerie médicale pour le troisième cité.
Il s'agit dès lors de faire comprendre aux systèmes informatiques et automatiques la forme telle que nous *imaginons* qu'elle est perçue par l'oeil humain. Nous pouvons assez bien imaginer qu'une forme est délimitée par un contour assez régulier, assez contrasté, séparant plusieurs régions présentant une certaine *homogénéité* à définir mathématiquement. Ce sont toutes ces propriétés intuitivement raisonnables — formalisées par les psychophysiciens (voir les travaux de Kanisza [Ka]) — que l'on cherche à incorporer dans des modèles mathématiques et que l'on implémente ensuite sur des systèmes informatiques.

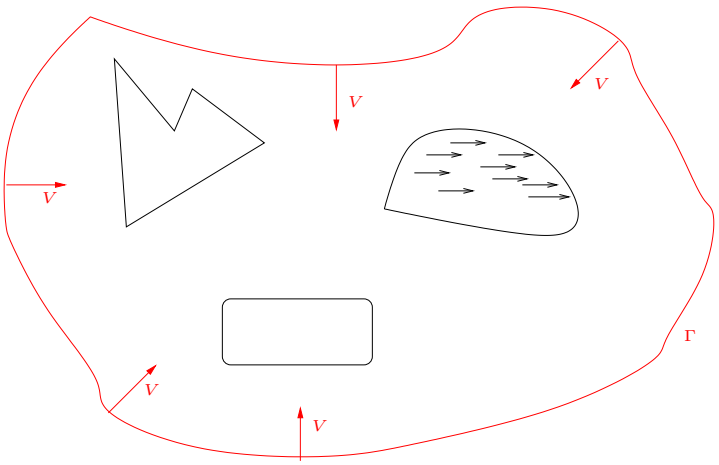


FIG. 1.1: Illustration du principe des contours actifs : on fait évoluer la courbe Γ par le champ de vecteurs V jusqu'à segmenter les objets à segmenter.

1.2 Problématiques des contours actifs

Pour résoudre le problème de détection de bords ou de segmentation d'une image (problèmes duaux l'un de l'autre), certains ont imaginé ([KWT], [CCCD], [CKS], [Co]...) utiliser des *contours actifs*. On se donne une formalisation continue du problème, c'est-à-dire que l'image initialement définie sur une grille rectangulaire est prolongée sur l'ouvert $D \subset \mathbb{R}^2$ correspondant. On remplace donc la suite finie $I_{k,l}$ des valeurs de l'image prises sur la grille par une fonction $I : D \to \mathbb{R}$. L'idée d'interpréter les images et les différents éléments qui la constitue comme des versions discrétisées de grandeurs mathématiques continues a longtemps prouvée son efficacité.
Le principe des contours actifs est de se donner une courbe (*cf.* **figure 1.1**) et de la faire évoluer de façon à ce qu'elle atteigne un état stable et qu'elle satisfasse au but que l'on se serait fixé (segmenter les différents objets de la scène, les objets répondant à un certain critère géométrique ou statistique, ou encore les objets en mouvement...). De manière générale, l'évolution de cette courbe se fait en minimisant un critère qui exprime certaines contraintes. On peut faire intervenir des contraintes sur l'aire, le périmètre, la courbure ou d'autres grandeurs géométriques relatives à la courbe. On peut aussi vouloir que les statistiques de l'image ou plus généralement des données suivent une certaine loi de distribution sur chacune des deux régions délimitées par la courbe. Les premiers modèles de contours actifs ne faisaient intervenir que des intégrales de contours du type

$$\int_{\partial\Omega} k^b \, d\mathcal{H}^1.$$

Ces dernières années, des modèles dits "basés régions" ont été de plus en plus utilisés. La différence est que l'on cherche à minimiser plusieurs critères sur les régions délimitées par le contour, c'est-à-dire que la fonctionnelle à minimiser est du type

$$\int_{\Omega} k^{in}\, d\mathbf{x} + \int_{\Omega^c} k^{out}\, d\mathbf{x} + \int_{\partial\Omega} k^{b}\, d\mathcal{H}^1.$$

Nous détaillerons dans le premier chapitre la variété des applications d'un tel modèle générique dû à Aubert *et al.* ([ABJB2]). L'un des aspects intéressants d'un tel modèle est de ne pas s'interdir que les fonctions k^{in} ou k^{out} dépendent elles-mêmes de la forme Ω ou que k^b dépende des propriétés de $\partial\Omega$. Ce modèle très générique recouvre la plupart des modèles de contours actifs basés régions existants ([CV1], [Ro], [ZY], [DP1],...).
La minimisation du critère se fait usuellement par une descente de gradient et donc par une EDP.
Il existe aussi des algorithmes pour minimiser le critère dans sa version discrétisée sur la grille de l'image. La fonctionnelle est réécrite en termes discrets, on se donne une fonction binaire $\phi \in \{-1, +1\}$ telle que $\Omega = \{\phi = 1\}$ (on note s un pixel, $l(s)$ la discrétisation de la mesure continue $d\mathcal{H}^1$ et H la fonction de Heaviside), ce qui donne

$$\sum_s k^{in} H(\phi(s)) + \sum_s k^{out}(1 - H(\phi(s))) + \sum_s k^b\, l(s).$$

Les algorithmes de minimisation de cette énergie se justifient dans le cadre d'une modélisation *markovienne* de la carte de la segmentation, c'est-à-dire que la valeur prise en un pixel s ne dépend que des valeurs prises en les pixels voisins, ce qui en terme probabilistes s'écrit (on note Φ le champ de variable aléatoire dont ϕ est la réalisation et $\mathcal{V}$ un *système de voisinage* [on précisera le sens mathématique dans le chapitre 2], ϕ_{-s} désigne les valeurs de ϕ en tous les pixels différents de s)

$$\mathbb{P}(\Phi_s = \phi_s |\, \Phi_{-s} = \phi_{-s}) = \mathbb{P}(\Phi_s = \phi_s |\, \Phi_{s'} = \phi_{s'},\ s' \in \mathcal{V}(s)).$$

Dans certains cas, il peut être plus avantageux d'utiliser de tels algorithmes qui procèdent par mise à jour successive des pixels en figeant les valeurs prises sur les autres pixels.
Les premiers modèles de restauration et de segmentation d'images dus à Geman et Geman [Ge] étaient des modèles discrets. Mumford et Shah [MS2] s'en étaient inspirés pour proposer une fontionnelle continue qui est à l'origine de la plupart des travaux portant sur la segmentation d'images en régions homogènes. Des algorithmes coûteux en temps de calcul comme le recuit simulé ont été proposés pour minimiser les énergies discrètes dues à Geman et Geman. L'algorithme que nous allons utiliser dans cette thèse est l'Iterated Conditional Modes (ICM) beaucoup plus rapide mais *a priori* plus sensible aux minima locaux que les algorithmes par EDP dérivés de la formulation continue du problème.

1.3 Problématiques liées à la modélisation du mouvement

Lorsque l'on a affaire à une vidéo, nombre de difficultés liées au mouvement apparaissent, complexifiant un peu plus l'analyse par rapport aux difficultés précitées concernant l'interprétation d'une image. La première difficulté est due au fait que l'image n'est que la projection d'une scène 3D sur le plan de la caméra, et que donc le mouvement *apparent* est différent du mouvement 3D réel des objets. L'autre difficulté tient à la modélisation mathématique du problème de la détermination du mouvement apparent ou *flot optique* ; bien souvent, on est conduit à résoudre un problème mal posé. En effet, en faisant l'hypothèse de conservation de l'intensité, le flot optique $\mathbf{v}$ doit satisfaire à l'équation

$$< \nabla I, \mathbf{v} > + \partial_t I = 0 \tag{1.1}$$

où ∇I désigne le gradient de l'image et $\partial_t I$ la dérivée partielle par rapport au temps de I. Cette équation permet donc de déterminer la composante de $\mathbf{v}$ suivant ∇I (le flot normal) mais l'autre composante reste indéterminée : c'est le *problème d'ouverture* (*cf.* **figure 1.2**). La première solution habituellement envisagée est de rajouter des contraintes de régularité et de prendre pour $\mathbf{v}$ la solution d'un problème de minimisation du type

$$\inf \left\{ \int_D | < \nabla I, \mathbf{v} > + \partial_t I|^2 \, d\mathbf{x} + \alpha R(\mathbf{v}) \right\}.$$

La deuxième solution est de supposer que le flot optique est constant sur un voisinage $\mathcal{V}$ de taille fixe autour de chaque pixel et d'intégrer l'équation du flot optique sur ce voisinage. Là encore, $\mathbf{v}$ est choisi comme la solution d'un problème de minimisation

$$\inf \left\{ \int_{\mathcal{V}} | < \nabla I, \mathbf{v} > + \partial_t I|^2 \, d\mathbf{x} \right\},$$

ce qui donne une expression assez simple pour le flot optique puisqu'on le suppose constant sur $\mathcal{V}$ (*cf.* [LK] et [BGW]).
Une autre difficulté est liée à la discrétisation des opérateurs ∇ et ∂_t sur la grille de l'image. Il s'agit du problème dit de *localité* du flot optique. Supposons que le flot optique *réel* soit d'une norme beaucoup plus grande que la taille du pixel que nous prendrons égale à 1. La façon la plus naturelle et la plus usuelle de discrétiser les opérateurs de dérivées est de les approcher par des différences finies. On a donc

$$(\partial_x I)_{i,j,k} \approx I_{i+1,j,k} - I_{i-1,j,k} \quad (\partial_y I)_{i,j,k} \approx I_{i,j+1,k} - I_{i,j-1,k},$$
$$(\partial_t I)_{i,j,k} \approx I_{i,j,k+1} - I_{i,j,k}.$$

On essaye donc d'estimer le flot optique à partir des valeurs de I sur un voisinage de taille beaucoup plus petit que la norme du flot optique théorique. On

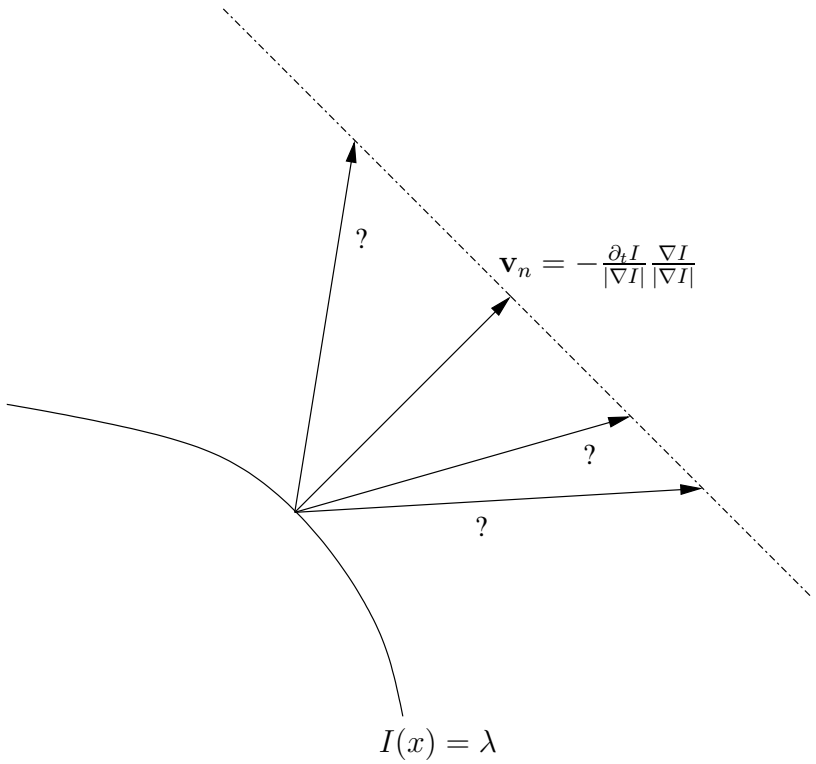


FIG. 1.2: Illustration du problème d'ouverture : sur une ligne de niveau de l'image (l'ensemble $\{\mathbf{x}|, I(\mathbf{x}) = \lambda\}$), seule la composante normale du flot est calculable. Les autres vecteurs de même composante normale sont également des valeurs possibles du flot optique, mais l'équation de contrainte ne permet pas de déterminer lequel choisir pour le flot optique.

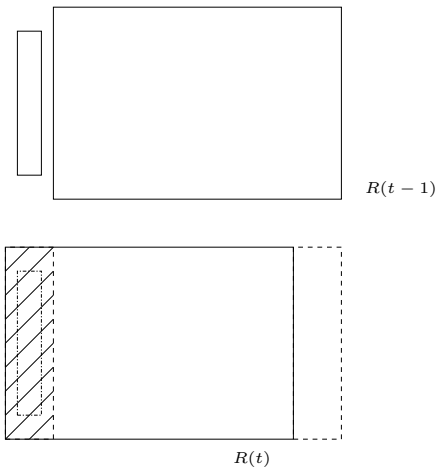


FIG. 1.3: Le rectangle est dans la position $R(t-1)$ sur l'image $I(\cdot, t-1)$ et $R(t)$ (rectangle en traits pleins, la position précédente étant en traits pointillés) sur l'image $I(\cdot, t)$.

voit bien que cela a peu de chance d'aboutir, car pour l'estimer correctement, il faudrait que le voisinage qui sert à la discrétisation soit d'une taille comparable à la norme du flot réel. Il semble donc quasiment impossible d'avoir une bonne estimation des grands déplacements à partir des techniques standards utilisant l'équation de contrainte du flot optique.
La première solution qui a été proposée pour circonvenir ce problème est de ne pas linéariser l'équation de conservation de l'intensité. Autrement dit, la conservation de l'intensité s'exprime simplement par $I(\mathbf{x}, t) = I(\mathbf{x} + \mathbf{v}, t + 1)$ et non plus par l'équation 1.1 faisant intervenir les dérivées spatiales et temporelles de l'image. On n'est donc plus limité par le problème de discrétisation. On se reportera à [NE] ou à [AWS] pour des exemples de cette idée. La deuxième solution est assez proche de celle-ci puisqu'elle consiste à linéariser l'équation de conservation du flot optique, mais progressivement, en partant d'une version sous-échantillonnée de l'image et en cherchant l'incrément autour de la valeur trouvée à cette résolution. Pour cette raison, de telles méthodes sont dénommées méthodes *multirésolution* (voir [MP])
Enfin, la dernière difficulté est dûe aux *occultations*. Considérons le diagramme suivant (*cf.* **figure 1.3**) d'un rectangle en mouvement. La partie hachurée sur l'image $I(t)$ recouvre un objet présent à cet endroit (le petit rectangle en pointillé) dans l'image $I(t-1)$, et donc la conservation de l'intensité n'est plus vérifiée sur cette région. Dans cette thèse, nous ne considérerons pas le problème des occultations. On pourra se reporter à [Sa] pour une étude du phénomène d'occultation dans les séquences d'images.

1.4 Solutions proposées

1.4.1 Segmentation et tracking d'objets en mouvement

1.4.1.1 Utilisation du flot optique

Une étude de plusieurs modèles de flot optique montre que l'un des modèles les plus performants pour la détermination du flot optique est celui de Weickert et Schnörr. Il s'agit d'un modèle variationnel où le flot optique est supposé avoir une régularité en espace et aussi en temps, donc la fonctionnelle à minimiser est de la forme (∇^θ désigne le gradient spatio-temporel (∇, ∂_t))

$$\int_{D\times[0,T]} | < \nabla I, \mathbf{v} > +\partial_t I|^2 \, d\mathbf{x}dt + \alpha \int_{D\times[0,T]} \Psi(|\nabla^\theta v_1|^2 + |\nabla^\theta v_2|^2) \, d\mathbf{x}dt.$$

Weickert et Schnörr ont montré dans [WS] que l'hypothèse de régularité spatio-temporelle du flot optique permettait de réduire les effets de bruit par rapport à une régularité spatiale. Ils ont également comparé leur algorithme sur des séquences dont le mouvement réel est connu et ont obtenu des résultats meilleurs que pour d'autres algorithmes en terme d'erreur angulaire entre le flot optique obtenu et le flot optique réel. Nos expériences confirment que cet algorithme est bien l'un des meilleurs à l'heure actuelle (*cf.* **figures 1.4 et 1.5**).
Nous avons parfois trouvé utile de combiner l'algorithme de Weickert et Schnörr avec une procédure multirésolution de compensation des grand mouvements. En effet, un autre avantage de ce genre de procédure est de mieux remplir les zones homogènes de l'image : lorsque certains objets en mouvement comportent de grandes zones très homogènes en niveau de gris, le flot optique est généralement mal déterminé sur ces zones car l'équation (1.1) ne nous fournit aucune information sur le flot optique dans ce cas.
Dans la procédure multirésolution, on cherche le flot optique entre les images sous-échantillonnées, les zones homogènes sont donc réduites et leur effet sur le flot optique moindre. Lorsque l'on répercute l'estimation du flot optique à la résolution supérieure, on "remplit" les zones homogènes avec un flot optique de norme plus fiable que si on le cherchait directement à cette résolution.
Une fois réglé le problème d'estimation du flot optique, on cherche à déterminer les objets en mouvement. L'idée simple sur laquelle nous nous sommes basés est de seuiller la norme du flot optique. On suppose donc que celle-ci est assez grande sur les objets en mouvement. Cette hypothèse est relativement bien vérifiée pour la plupart des méthodes de flot optique utilisées (*cf.* **figures 1.4 et 1.5**) mais la méthode de Weickert et Schnörr fournit un résultat plus précis. Toutefois, le résultat donné par un calcul de flot optique comporte toujours du bruit et de plus un seuillage de la norme du flot ne donne pas exactement les contours des objets. Afin de résoudre ces problèmes, il nous a paru intéressant de considérer le problème d'optimisation de forme suivant : on cherche le domaine Ω (représentant l'ensemble des objets en mouvement) tel que le flot optique soit

FIG. 1.4: Une image extraite d'une séquence. Les différents objets en mouvement sont les voitures.

de norme plus grande qu'un certain seuil α sur Ω, que la frontière de Ω soit suffisamment régulière, et aussi que le gradient de l'image soit suffisamment fort le long de cette frontière, ce qui s'exprime par la minimisation de la fonctionnelle

$$E(\Omega) = \int_{\Omega} \alpha \, d\mathbf{x} + \int_{D\setminus\Omega} |\mathbf{v}| \, d\mathbf{x} + \nu \int_{\partial\Omega} d\mathcal{H}^1(\mathbf{x}) + \lambda \int_{\partial\Omega} \underbrace{\frac{1}{1+|\nabla(G_\sigma \star I)(\mathbf{x})|^2}}_{g_I(\mathbf{x})} \, d\mathcal{H}^1(\mathbf{x}).$$

1.4.1.2 Optimisation de forme et calcul des variations

La dérivation d'une telle fonctionnelle est un problème non trivial en soi. Il existe deux méthodes majeures :

– La première façon consiste à conserver la fonctionnelle et à définir une nouvelle dérivée qui tienne compte d'une pertubation Ω_t de la forme Ω par un champ de vecteurs V. La *dérivée eulérienne* de E est définie alors comme

$$dE(\Omega; V) = \lim_{t\to 0} \frac{E(\Omega_t) - E(\Omega)}{t}.$$

  Nous présenterons en annexe l'ensemble des résultats d'optimisation de forme et des calculs de dérivées eulériennes utiles dans cette thèse.

– La deuxième façon est d'utiliser la technique dite "des fonctions de Heaviside" qui consiste à remplacer Ω par une fonction u positive sur Ω et négative sur $D \setminus \Omega$. On peut alors représenter Ω comme l'ensemble $\{\mathbf{x}|\, u(\mathbf{x}) > 0\} =$

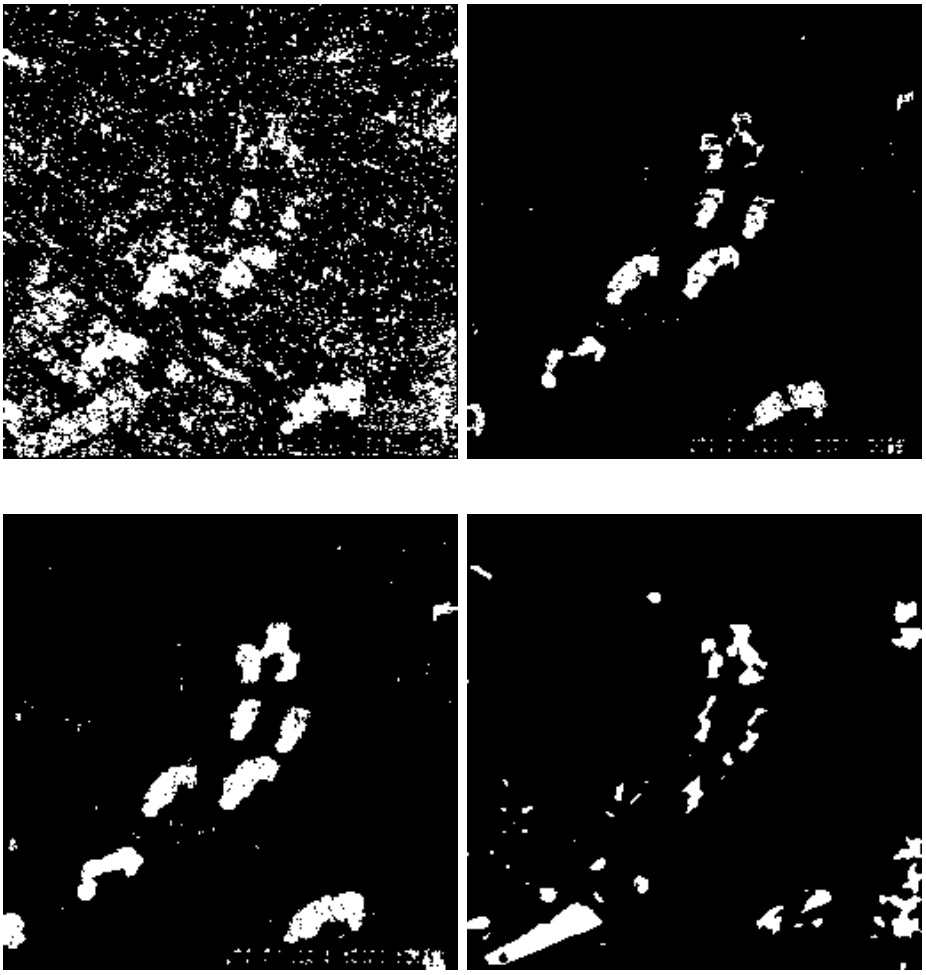


FIG. 1.5: Les masques montrés sont obtenus en seuillant la norme du flot optique sur l'image précédente. De gauche à droite et de haut en bas : Horn et Schunck, Weickert-Schnörr, Weickert-Schnörr avec multirésolution et la méthode des tenseurs de Bigün. Pour les trois premières, le flot est seuillé à $0,7$, pour la dernière à $0,5$ (visuellement le meilleur). Les résultats les plus convaincants sont obtenus pour le modèle de Weickert et Schnörr, que ce soit avec ou sans multirésolution.

$\{\mathbf{x}|\, H(u(\mathbf{x})) = 1\}$ et remplacer l'énergie E par une énergie fonction de u

$$J_\epsilon(u) = \int_D \alpha H_\epsilon(u(\mathbf{x}))\, d\mathbf{x} + \int_D |\mathbf{v}|(\mathbf{x})(1 - H_\epsilon(u(\mathbf{x})))\, d\mathbf{x} + \int_D (\lambda g_I(\mathbf{x}) + \nu)|\nabla H_\epsilon(u(\mathbf{x}))|\, d\mathbf{x}.$$

On voit apparaître un paramètre ϵ qui sert à construire une version régularisée H_ϵ de la fonction de Heaviside. Cette étape est nécessaire pour pouvoir calculer la dérivée de Gâteaux de la fonctionnelle qui ne pourrait se faire avec la fonction de Heaviside.

La minimisation se fait par descente de gradient : on cherche V tel que $dE(\Omega; V) < 0$ ou w tel que $dJ(u; w) < 0$. En utilisant la notion de dérivées de forme, on obtient une équation aux dérivées partielles sur la frontière Γ de Ω (on note κ la courbure de Γ et $\mathbf{n}$ sa normale extérieure)

$$\frac{\partial \Gamma}{\partial t} = (|\mathbf{v}| - \alpha + (\lambda g_I + \nu)\kappa + \lambda < \nabla g_I, \mathbf{n} >)\, \mathbf{n}.$$

Avec la technique "des fonctions de Heaviside", on obtient l'EDP

$$\frac{\partial u}{\partial t} = \delta_\epsilon(u) \left(|\mathbf{v}| - \alpha + \operatorname{div} \left((\lambda g_I + \nu) \frac{\nabla u}{|\nabla u|} \right) \right).$$

Mais l'important est que les deux approches envisagées (dérivée eulérienne ou fonctions de Heaviside) conduisent aux mêmes mouvements de courbe.

Le temps de calcul d'une telle EDP implémentée avec un schéma semi-implicite AOS (*cf.* [RVW]) sur un PC doté d'un processeur Pentium IV cadencé à 1.8 GHz et de 1 Go de RAM est de l'ordre de 1 ou 2 minutes, ce qui est très long si on le compare avec les temps de calcul d'algorithmes discrets comme l'ICM (voir plus loin), mais relativement raisonnable. De plus, l'utilisation de l'ICM sur un tel modèle faisant intervenir un terme de bord non constant ($\lambda g_I + \nu$ ici) est moins évident à cause d'effets de pixelisation déjà assez marqués dans le cas d'un terme de bord constant (cela sera illustré dans le chapitre 4) et qui le sont encore plus fortement lorsque l'on introduit une fonction dépendant de la position dans le terme de bord. Pour cette raison nous n'avons pas appliqué l'ICM à ce modèle.

La méthode que nous avons proposée est donc une méthode d'amélioration de la qualité du seuillage de la norme du flot optique. Elle est particulièrement indiquée lorsque les objets ont un mouvement rigide, bien que nous ayons aussi obtenu d'assez bons résultats sur des séquences d'images avec des mouvements articulés.

La généralisation de l'algorithme au cas où la caméra n'est pas fixe se fait en calculant le mouvement induit par le déplacement de caméra et en effectuant un *recalage* (suppression virtuelle du mouvement de caméra) des images. On traite alors la séquence d'images recalées.

1.4.1.3 Application de l'ICM à une classe de problèmes d'ensembles de niveaux

L'autre étude principale de cette thèse porte sur des modèles qui sont caractérisés par la segmentation de l'image en deux régions Ω et $D \setminus \Omega$ et où Ω doit minimiser une fonctionnelle du type

$$E(\Omega) = \int_{\Omega} f_1(\mathbf{x})\, d\mathbf{x} + \int_{D\setminus\Omega} f_2(\mathbf{x})\, d\mathbf{x} + \nu \int_{\partial\Omega} d\mathcal{H}^1(\mathbf{x}). \tag{1.2}$$

Nous avons vu que dans sa version discrète, cette fonctionnelle s'écrit

$$\sum_s f_1(s)H(\phi(s)) + \sum_s f_2(s)(1 - H(\phi(s))) + \nu \sum_s l(s).$$

L'idée de discrétiser la fonctionnelle continue pour la minimiser dans un cadre discret a été récemment réappliquée par Chan et Song dans [SC] à la fonctionnelle de Mumford et Shah dans le cas constant par morceaux ([MS2]). Chan et Song n'avaient toutefois pas noté l'équivalence de leur algorithme avec l'ICM. Les premiers travaux sur l'équivalence entre modèles discrets et continus sont dus à Chambolle qui démontre dans [Cham] la Γ-convergence de la solution du problème discret de Geman et Geman vers la solution d'une fonctionnelle qui est une version anisotropique de la fonctionnelle de Mumford et Shah ([MS1]). Enfin, le terme de bord $\sum_s l(s)$ se voit être la discrétisation exacte du terme de longueur. Certaines fonctionnelles discrètes utilisent des termes de bords qui ne correspondent pas à la longueur en description continue. Ainsi, nous verrons que les termes de bord basés sur le modèle d'Ising sont issus de la discrétisation d'une longueur "pondérée" mais pas de la longueur au sens usuel. Nous montrerons que la minimisation de la fonctionnelle discrète peut s'implémenter efficacement par l'algorithme des Modes Conditionnels Itérés, en anglais Iterated Conditional Modes (ICM). Cet algorithme présente l'intérêt d'être beaucoup plus rapide qu'un algorithme par EDP. En effet, le temps de calcul de l'implémentation de l'EDP de descente de gradient de l'énergie

$$\frac{\partial u}{\partial t} = \delta_\epsilon(u) \left(f_2(\mathbf{x}) - f_1(\mathbf{x}) + \nu \mathrm{div} \left(\frac{\nabla u}{|\nabla u|} \right) \right)$$

est de l'ordre de 1 minute alors qu'il est de l'ordre de quelques secondes pour l'ICM. Enfin, nous verrons que l'utilisation des EDP reste avantageuse à certains points de vue, car elle évite certains désagréments (irrégularité, aliasing du contour) dus à l'effet très local (la valeur de la carte de segmentation ϕ est mise à jour pixel par pixel en gelant les valeurs en tous les autres pixels) de l'ICM.

1.4.2 Problèmes de vidéo surveillance

Dans la dernière partie de cette thèse, nous nous attacherons à deux problèmes concrets de vidéo surveillance. Le premier est de pouvoir estimer la densité d'une

foule sur un quai de métro, le deuxième de pouvoir détecter des comportements "anormaux" de personnes dans l'environnement du métro parisien.
Le problème d'estimation de la densité d'une foule étant davantage un problème d'image qu'un problème de vidéo, nous avons choisi de faire des calculs de courbure sur chaque image, en faisant l'hypothèse assez raisonnable que celle-ci devrait être forte lorsque la foule est très dense et faible lorsque le quai de métro est vide (il y a soit des zones homogènes dans l'image, soit des éléments généralement rectilignes).
Concernant la détection des comportements anormaux, l'idée est de calculer un flot optique et d'étudier sa direction. Comme les contraintes sont ici plus sévères (rapidité, voire temps réel), l'utilisation d'algorithmes de flot optique tels que celui de Weickert et Schnörr devient difficile car elle implique un coût de calcul assez élevé, à moins d'utiliser des stratégies multigrilles (voir des travaux récents de J. Weickert et ses collaborateurs dans le cas de la fonctionnelle de Horn et Schunck [BWFKS]). Pour cette raison, nous avons plutôt utilisé des méthodes plus simples (Lucas et Kanade [LK] ou Bernard [Ber]) où le flot est supposé constant sur un voisinage de chaque pixel. Les mouvements dits "anormaux" sont ceux où on repère des directions du flot optique inhabituellement regroupées autour de directions préalablement définies comme correspondant à l'anormalité.

1.5 Plan du mémoire

- La première partie de cette thèse (chapitres 2, 3 et 4) porte sur la segmentation et le tracking d'objets en mouvement dans une vidéo.
  - Dans le chapitre 2, nous rappelons les techniques relatives aux contours actifs ainsi que différents modèles généraux qui ont été proposés pour résoudre le problème de la segmentation d'images fixes. En effet, un certain nombre de modèles de segmentation vidéo s'inspirent de travaux qui ont été faits sur la segmentation d'images. Il s'agit donc d'un état de l'art sur la segmentation d'images.
  - Les chapitres 3 et 4 sont relatifs à la segmentation et au tracking d'objets en mouvement pour lesquels nous avons proposé les solutions exposées dans la section précédente. Le chapitre 3 est consacré à ces problèmes dans le cas où la caméra est fixe et le chapitre 4 traite du cas où la caméra est en mouvement. Le chapitre 3 inclut également un bref état de l'art sur les travaux importants sur ce sujet : segmentation conjointe fond/objets en mouvement (section 3.1) et revue des méthodes de flot optique et des méthodes de segmentation par flot optique (section 3.3).
- La deuxième partie de cette thèse (chapitre 5) porte sur l'étude des deux problèmes de vidéo surveillance cités plus haut et qui nous ont été soumis dans le cadre d'un contrat avec la Régie Autonome des Transports Parisiens (RATP).

– L'avant-dernier chapitre est un chapitre d'annexe qui regroupe l'ensemble des résultats d'optimisation de forme utiles dans cette thèse. On y trouvera un récapitulatif des résultats importants fréquemment utilisés dans les autres chapitres. A chaque fois que nous utilisons un de ces résultats dans un des autres chapitres, il est fait référence à cette annexe. Le calcul de *dérivées de domaine* a été introduit par Sokolowski et Zolésio dans [SZ] et constitue une nouvelle approche pour l'optimisation de forme. La plupart des fonctionnelles utilisées dans les modèles de contours actifs peuvent se dériver en utilisant les dérivées de domaine. Pour cette raison, nous avons regroupé les principales définitions et les principaux résultats dans un seul chapitre.
– Les algorithmes que nous avons utilisé ont été implémenté en MegaWave, environnement de programmation libre développé par J. Froment et maintenu par J. Froment et L. Moisan. Le dernier chapitre regroupe les documentations relatives aux modules développés pour l'implémentation des algorithmes.

Chapitre 2

Méthodes de segmentation d'images fixes

Ce chapitre est consacré à la segmentation d'images par contours actifs hors du contexte de la vidéo. Il s'agit donc de détecter les principaux objets dans une scène fixe. Notre but est de présenter les différents modèles de contours actifs, leurs motivations et les résultats correspondants afin de nous en inspirer pour la segmentation vidéo exposée dans les chapitres 3 et 4. Il est à signaler que nous présenterons des méthodes statistiques discrètes de segmentation plus loin dans le chapitre 3.

2.1 Présentation des contours actifs

2.1.1 Le Principe

On présente ici le principe des contours actifs ou *snakes* sur un domaine $D \subset \mathbb{R}^2$. Bien que ce principe ait été généralisé à un volume $D \subset \mathbb{R}^3$ puis à des variétés, on ne traite pas de cela ici.
Une image I étant donnée sur un ouvert $D \subset \mathbb{R}^2$, il s'agit de faire évoluer une courbe paramétrée $\Gamma_0 : [0, L(\Gamma_0)] \to \mathbb{R}^2,\ s \mapsto (x(s), y(s))$ et de la faire converger vers les bords des objets qui nous intéressent. On peut alors soit introduire un potentiel caractéristique de ces bords (qui soit idéalement nul sur les bords et non nul ailleurs) ([KWT], [CCCD], [CKS], [Co]), soit dire que la courbe est une interface entre deux régions (son intérieur et son extérieur) qui doit atteindre un certain équilibre dépendant des propriétés de ces deux régions, en n'excluant pas de se servir également d'un potentiel caractéristique des bords à atteindre. Cette dernière approche est classiquement appelée "contours actifs basés régions"et a donnée lieu à une vaste littérature ([CV1], [ABJB2], [DP2], [SBFAZ] pour ne citer que quelques références qui donnent un panorama des champs d'application de cette approche).
Initialement introduits par Kass, Witkin et Terzopoulos dans le cadre d'une

formulation variationnelle ([KWT]), les contours actifs ont été adaptés par la suite par Caselles, Catté, Coll et Dibos dans [CCCD] où les auteurs écrivent une équation aux dérivées partielles proche de l'équation de la courbure moyenne et qui ne peut être obtenue par minimisation d'une énergie. Plus tard, dans [CKS], Caselles, Kimmel et Sapiro ont adopté une approche variationnelle moins contraignante que celle de Kass, Witkin et Terzopoulos. On ne présente pas ces trois méthodes dans l'ordre chronologique décrit ci-dessus, mais plutôt dans l'ordre Kass-Witkin-Terzopoulos, Caselles-Kimmel-Sapiro, Caselles-Catté-Coll-Dibos, les deux premières ayant en commun d'être des méthodes variationnelles.

2.1.2 Le premier modèle variationnel de Kass, Witkin et Terzopoulos

Dans le cadre variationnel, le problème est de rechercher la courbe paramétrée $\Gamma : [a, b] \to \mathbb{R}$ avec $\Gamma(a) = \Gamma(b)$ (courbe fermée) qui minimise une énergie $J(\Gamma)$. Dans le cas de la fonctionnelle de Kass, Witkin et Terzopoulos, elle s'écrit

$$J_{KWT}(\Gamma) = \underbrace{\int_a^b g(|\nabla(G_\sigma \star I(\Gamma(s)))|)^2 \, ds}_{terme\ d'attraction} + \underbrace{\int_a^b (\alpha \, |\Gamma'(s)|^2 + \beta |\Gamma''(s)|^2) \, ds}_{terme\ de\ régularisation}$$

où ∇I désigne le gradient de l'image I et α et β des paramètres positifs. Initialement, Kass, Witkin et Terzopoulos choisirent $g(|\nabla(G_\sigma \star I)|) = -\lambda |\nabla(G_\sigma \star I)|^2$ mais il ne pose aucun problème de généraliser cette formulation.
Le premier terme de cette fonctionnelle tend à attirer la courbe vers les bords. Le potentiel est donné par la fonction $g(|\nabla(G_\sigma \star I)|)$ où g est une fonction décroissante qui tend vers 0 en $+\infty$, mais en général, on prend

$$g(|\nabla(G_\sigma \star I)|) = \frac{1}{1 + |\nabla(G_\sigma \star I)|^2} \text{ avec } G_\sigma(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{x^2 + y^2}{2\sigma^2})$$

On peut éventuellement rajouter un paramètre d'échelle δ visant à couper les gradients inférieurs à ce paramètre

$$g(|\nabla(G_\sigma \star I)|) = \frac{1}{1 + |\frac{\nabla(G_\sigma \star I)}{\delta}|^2}$$

En un bord idéal, on a $|\nabla I| = +\infty$ et donc le potentiel et le terme d'attraction valent 0, donc la minimisation de ce terme tend bien à attirer la courbe vers les bords.
Le deuxième terme est un terme qui impose une certaine régularité à la courbe, notamment que la courbure soit assez faible. Comme on minimise la somme de ces deux termes, en pratique, on est amené à trouver un compromis entre régularisation et convergence vers les bords.

L'inconvénient de ce modèle était de dépendre de la paramétrisation de la courbe Γ et également d'empêcher que celle-ci se scinde en plusieurs courbes au cours de son évolution. Un tel cas induirait une explosion de la courbure, mais celle-ci est précisément contrôlée par le terme de régularisation. Cette situation n'est pas très commode car on a alors besoin d'autant de courbes que d'objets à segmenter.

2.1.3 Le modèle variationnel de Caselles, Kimmel et Sapiro

Afin de surmonter les inconvénients décrits plus haut, Caselles, Kimmel et Sapiro choisirent de ne garder qu'un terme d'attraction de la forme

$$J_{CKS^1}(\Gamma) = \int_a^b g(|\nabla(G_\sigma \star I(\Gamma(s)))|)|\Gamma'(s)|\, ds$$

Cette fois, le modèle ne dépend plus de la paramétrisation de la courbe puisque l'on peut reparamétriser la courbe par son abscisse curviligne. Le terme $\Gamma''(s)$ a été supprimé car on peut montrer que le modèle de Kass, Witkin et Terzopoulos avec $\beta = 0$ diminue lui aussi la courbure, il s'agit donc d'un terme redondant. La fonctionnelle J_{CKS^1} peut donc être réécrite

$$J_{CKS^1}(\Gamma) = \int_\Gamma g(|\nabla(G_\sigma \star I(\mathbf{x}))|)\, d\mathcal{H}^1(\mathbf{x})$$

où $d\mathcal{H}^1(\mathbf{x})$ désigne la mesure de Hausdorff. On voit que la fonctionnelle peut être vue comme une longueur "pondérée" de la courbe par $g(|\nabla(G_\sigma \star I)|)$, ce qui revient à dire que nous définissons une nouvelle métrique et que nous cherchons les géodésiques, d'où l'appellation "contours actifs *géodésiques*". Implémentée par la méthode des ensembles de niveaux (*cf.* section 2.1.4.2), la minimisation de la fonctionnelle permet les changements de topologie non autorisés lorsque l'on conserve cette représentation non implicite de la courbe.
La minimisation de cette fonctionnelle est obtenue par une méthode de descente de gradient : la courbe Γ suit l'équation d'évolution ($\mathbf{n}$ est la normale extérieure à la courbe Γ et κ la courbure de Γ)

$$\begin{aligned} \frac{\partial \Gamma}{\partial t} &= -\left(g(|\nabla(G_\sigma \star I)|)\kappa + (\nabla(g(|\nabla(G_\sigma \star I)|)) \cdot \mathbf{n})\right)\mathbf{n} \\ \Gamma(t=0) &= \Gamma_0. \end{aligned}$$

Le premier terme du membre de droite tend à attirer la courbe vers un point, mais lorsque l'on atteint un bord idéal, on a $g = 0$ et donc l'influence de ce terme devient nulle.
Le deuxième terme est un terme d'érosion qui tend à attirer la courbe vers les bords des objets. Si $r^{-1}g'(r)$ admet une limite l quand $r \to 0$, il est de contribution nulle sur des zones de gradient nul de l'image, en effet, d'après les

théorèmes classiques de dérivation de fonctions composées,

$$\begin{aligned}\nabla(g(|\nabla I|)) &= g'(|\nabla I|)\nabla(\nabla I)\frac{\nabla I}{|\nabla I|}\\ &= \underbrace{|\nabla I|^{-1}g'(|\nabla I|)}_{\to l} H_I \underbrace{\nabla I}_{\to(0,0)}\\ &\to 0 \text{ quand } \nabla I \to (0,0)\end{aligned}$$

où H_I désigne la matrice hessienne de I.
Il est à noter qu'il est difficile de segmenter les objets non convexes en prenant le modèle tel quel. En effet, sur une zone homogène, on a $|\nabla I| \approx 0$ donc $g(|\nabla I|) \approx 1$ et $\nabla g(|\nabla I|) \approx \mathbf{0}$: le terme $g\kappa$ est prépondérant et fait tendre la courbe non vers les bords mais vers l'enveloppe convexe des bords.
Pour éviter cet écueil, on ajoute généralement un terme à l'équation pour obtenir

$$\begin{aligned}\frac{\partial \Gamma}{\partial t} &= -\left(g(|\nabla(G_\sigma \star I)|)(\kappa+\nu) + (\nabla(g(|\nabla(G_\sigma \star I)|)) \cdot \mathbf{n})\right)\mathbf{n}\\ \Gamma(t=0) &= \Gamma_0.\end{aligned} \tag{2.1}$$

ce qui correspond à la minimisation de

$$J_{CKS^2}(\Gamma) = \int_\Gamma g(|\nabla(G_\sigma \star I)|)\, d\mathcal{H}^1 + \nu \int_\Omega g(|\nabla(G_\sigma \star I(\mathbf{x}))|)\, d\mathbf{x}$$

où Ω est le domaine intérieur à Γ.
On ne présente pas ici l'intégralité des méthodes variationnelles de contour actif, une revue des méthodes autres que celles présentées ici ayant été faite par Desolneux, Moisan et Morel dans [DMM].

2.1.4 Le modèle d'évolution de Caselles, Catté, Coll et Dibos

2.1.4.1 Formulation scalaire

Enfin, Caselles, Catté, Coll et Dibos proposèrent de prendre comme équation d'évolution

$$\begin{aligned}\frac{\partial u}{\partial t} &= g(|\nabla(G_\sigma \star I)|)|\nabla u|\left(\operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right)+\nu\right) \quad (t,x) \in [0,\infty[\times\mathbb{R}^2\\ u(t=0) &= u_0\end{aligned} \tag{2.2}$$

où u_0 est une version régularisée de $1-\chi_C$, χ_C dénotant la fonction caractéristique d'un ensemble C tel que $\partial C = \Gamma$.
Ce modèle satisfait aux exigences suivantes :
- Il permet de trouver les contours qui reproduisent au mieux les contours des objets à segmenter dans l'image.
- Les contours obtenus sont réguliers.
- Les singularités comme les angles ou les coins se retrouvent sur la courbe limite obtenue pour l'état asymptotique de l'équation d'évolution.

– L'algorithme nécessite des paramètres, mais en nombre limité : le pas de temps, le nombre d'itérations et le facteur d'échelle δ dans la fonction g.
– Comme pour le modèle de Caselles, Kimmel et Sapiro, les changements de topologie sont autorisés.
– La théorie des solutions de viscosité d'une équation de Hamilton-Jacobi garantit l'existence de solutions de l'EDP.

La formulation adoptée par Caselles, Catté, Coll et Dibos est différente des précédentes, puisqu'elle n'est plus exprimée en terme d'évolution de courbe mais en terme d'évolution d'une fonction u définie sur $[0,\infty[\times\mathbb{R}^2$. Comme nous n'avons jusqu'ici exprimé le problème qu'en terme d'évolution de courbe, nous allons relier ces deux formulations. Le théorème suivant exprime la dualité entre évolutions scalaire et géométriques.

2.1.4.2 Dualité entre évolutions scalaire et géométriques, méthode d'Osher-Sethian

Théorème 1 *Soit I un intervalle ouvert de $\mathbb{R}$, et $u(x,t) : \mathbb{R}^2 \times I \to \mathbb{R}$ une solution régulière (C^2) de*

$$\frac{\partial u}{\partial t} = |\nabla u| F(\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)) \tag{2.3}$$

telle que pour tout réel λ, on a

$$\forall t \in I, \forall (x,y) \in C(t) = \{(x,y), u(x,y,t) = \lambda\}\ |\nabla u|(x,y,t) \neq 0.$$

Alors $C(t)$ suit l'équation d'évolution

$$\frac{\partial C}{\partial t} = -F(\kappa)\mathbf{n} \tag{2.4}$$

où $\mathbf{n} = \frac{\nabla u(C,t)}{|\nabla u(C,t)|}$ est le vecteur normal extérieur à C et $\kappa = \mathrm{div}\left(\frac{\nabla u(C,t)}{|\nabla u(C,t)|}\right)$ sa courbure.

Notons que si une équation d'évolution géométrique dérive d'une énergie, ce n'est pas le cas *a priori* pour l'équation d'évolution duale. En effet, la minimisation de la fonctionnelle périmètre $J(\Gamma) = \int_\Gamma d\mathcal{H}^1 = \mathcal{H}^1(\Gamma)$ est réalisée en faisant évoluer Γ par $\dfrac{\partial \Gamma}{\partial t} = -\kappa\,\mathbf{n}$ mais l'équation duale $\dfrac{\partial u}{\partial t} = |\nabla u| \mathrm{div}\left(\dfrac{\nabla u}{|\nabla u|}\right)$ ne peut être obtenue par descente de gradient d'une énergie puisque le terme de droite ne peut se mettre sous la forme d'une divergence.
La dualité entre évolution scalaire et géométrique est également avantageuse d'un point de vue pratique. En effet, implémenter l'équation (2.4) n'est pas un problème évident compte tenu de la forme de l'équation, de la localité et/ou de la paramétrisation de la courbe. Mais ce théorème nous dit que pour faire évoluer la courbe Γ par une équation d'évolution, on peut introduire une fonction u qui

admet cette courbe comme courbe de niveau et la faire évoluer selon l'équation duale. L'état final de la courbe est ainsi obtenu à partir de l'état final de la fonction. Cette méthode a été introduite par Osher et Sethian dans [OS] et est connue sous le nom de méthode des ensembles de niveaux (Level Set Method). En général, on prend pour cette fonction la fonction distance signée à la courbe

$$\begin{aligned} d(x,y) &= \epsilon(x,y)\min\{(x',y') \in \Gamma, \|(x-x',y-y')\|\} \\ \epsilon(x,y) &= +1 \text{ à l'intérieur} \\ &\quad -1 \text{ à l'extérieur.} \end{aligned}$$

Néanmoins, si la fonction distance semble très naturelle, son utilisation induit une difficulté supplémentaire, car généralement elle n'est pas solution de l'EDP que l'on cherche à implémenter. Il y a là une contradiction entre la théorie et son implémentation. Pour surmonter cette difficulté, on choisit généralement de réinitialiser la fonction u à une fonction distance toutes les n itérations de (2.3) où n est à choisir par l'utilisateur. Néanmoins, on peut citer le travail de Faugeras et Gomes [FG] où les auteurs construisent une fonction qui reste une fonction distance au cours de son évolution et réconcilient ainsi la théorie et l'implémentation.

2.1.4.3 Formulation géométrique

Grâce à la dualité entre évolution scalaire et géométrique, on peut réécrire l'EDP (2.2) en terme d'évolution de courbe et on a

$$\frac{\partial \Gamma}{\partial t} = -g(|\nabla(G_\sigma \star I)|)(\kappa + \nu)\mathbf{n} \tag{2.5}$$

Il convient de comparer cette EDP à l'équation de la courbure moyenne

$$\frac{\partial \Gamma}{\partial t} = -\kappa\,\mathbf{n}. \tag{2.6}$$

Evans et Spruck avaient démontré ([ES]) que cette équation tendait à convexifier la courbe Γ, puis à la faire converger vers un cercle qui se réduit jusqu'à disparaître. Les auteurs de [CCCD] ont constaté expérimentalement que l'équation

$$\frac{\partial \Gamma}{\partial t} = -g(|\nabla(G_\sigma \star I)|)\kappa\,\mathbf{n}$$

tendait elle aussi à convexifier la courbe. Pour cette raison, ils ont rajouté un terme d'érosion afin de pouvoir détecter des objets non convexes.

2.2 Un modèle de contours actifs basés régions généralisant le modèle de Mumford et Shah

On présente dans cette section une formulation générique des contours actifs basés région adoptée par Aubert, Barlaud et Jehan-Besson dans

[ABJB2]. Ils donnent une signification physique à chacun des termes de leur fonctionnelle intégrés sur les régions : ce sont les **descripteurs** des régions. Une formulation encore plus générique peut être adoptée dans le cadre des contours actifs multiples et de la classification d'image, on se référera notamment à [CV2] et [SBFAZ].

2.2.1 Formulation locale

On reprend les mêmes notations que précédemment : D représente le domaine entier de l'image, on considère Ω un ouvert de D qui représente la région intérieure à Γ, Ω^c est donc la région extérieure. Dans le cas où Ω est la réunion de plusieurs composantes connexes, Γ est la réunion des frontières des différentes composantes connexes de Ω. Il s'agit là d'une convention que prennent les auteurs de [ABJB2] et nous verrons qu'une telle modélisation peut être gênante dans certains cas.
La formulation choisie par Aubert, Barlaud et Jehan Besson dans [ABJB2] est la suivante : on se donne trois fonctions $k^{(in)}(x,y,\Omega)$, $k^{(out)}(x,y,\Omega^c)$ et $k^{(b)}(x,y)$ qui sont les descripteurs de l'intérieur, de l'extérieur et du bord, et on écrit l'énergie J sous la forme

$$\begin{array}{rl} J(\Omega) & = \int_\Omega k^{(in)}(\mathbf{x},\Omega)\,d\mathbf{x} + \int_{\Omega^c} k^{(out)}(\mathbf{x},\Omega^c)\,d\mathbf{x} \\ & \quad + \int_\Gamma k^{(b)}\,d\mathcal{H}^1. \end{array} \tag{2.7}$$

La dérivation de ce critère donne (voir l'annexe A)

$$\begin{array}{rl} dJ(\Omega;V) = & \int_\Omega (k^{(in)})'(\Omega;V)\,d\mathbf{x} + \int_{\Omega^c} (k^{(out)})'(\Omega^c;V)\,d\mathbf{x} \\ & + \int_\Gamma (k^{(in)} - k^{(out)} + k^{(b)}\cdot\kappa + < \nabla k^{(b)}, \mathbf{n} >) < V, \mathbf{n} >\, d\mathcal{H}^1 \end{array}$$

Dans [ABJB2], Aubert, Barlaud et Jehan-Besson proposent plusieurs descripteurs dont entre autres des fonctions de la variance intérieure σ^2_Ω de I sur Ω et extérieure $\sigma^2_{\Omega^c}$ sur Ω^c. On donne la définition des moyennes et variances intérieures (*resp.* extérieures) de I

$$\begin{array}{rclcrcl} \mu_\Omega & = & \frac{\int_\Omega I(x,y)\,d\mathbf{x}}{|\Omega|} & & \mu_{\Omega^c} & = & \frac{\int_{\Omega^c} I(x,y)\,d\mathbf{x}}{|\Omega^c|} \\ \sigma^2_\Omega & = & \frac{\int_\Omega [I(x,y)-\mu_\Omega]^2\,d\mathbf{x}}{|\Omega|} & & \sigma^2_{\Omega^c} & = & \frac{\int_{\Omega^c} [I(x,y)-\mu_{\Omega^c}]^2\,d\mathbf{x}}{|\Omega^c|}. \end{array}$$

Les auteurs de [ABJB2] écrivent donc $k^{(in,out)} = \phi(\sigma^2_{\Omega,\Omega^c})$ avec ϕ choisi comme $\phi(r) = \log(1+r)$ ou $\phi(r) = \sqrt{\epsilon + r^2}$. Remarquons que le cas $\phi(r) = r$ et $k^{(b)} = cste = \lambda$ donne le modèle constant par morceaux de Mumford et Shah avec simplement deux régions

$$J(\Omega) = \int_\Omega (I-\mu_\Omega)^2\,d\mathbf{x} + \int_{\Omega^c} (I-\mu_{\Omega^c})^2\,d\mathbf{x} + \lambda \int_\Gamma d\mathcal{H}^1;$$

on peut donc considérer la fonctionnelle d'Aubert *et al.* comme une généralisation de ce modèle. Les fonctions $\phi(r) = \log(1+r)$ et $\phi(r) = \sqrt{\epsilon + r^2}$ ont la

propriété d'atténuer les grandes valeurs de r et donc d'être assez robustes à une éventuelle mesure erronée de la variance.
Dans ce cas on peut calculer (voir le chapitre consacré à l'optimisation)

$$\int_\Omega (k^{(in)})'(\Omega;V)\,d\mathbf{x} = 2\sigma_\Omega \sigma'_\Omega(\Omega;V)\phi'(\sigma^2_\Omega)|\Omega| = \phi'(\sigma^2_\Omega)\int_\Gamma [(I-\mu_\Omega)^2-\sigma^2] < V, \mathbf{n} > \, d\mathcal{H}^1$$

Ainsi, on diminue J en choisissant $V = -F\,\mathbf{n}$ où

$$\begin{aligned} F = \quad & \phi(\sigma^2_\Omega) + \phi'(\sigma^2_\Omega)((I-\mu_\Omega)^2-\sigma^2_\Omega) \\ & - \phi(\sigma^2_{\Omega^c}) - \phi'(\sigma^2_{\Omega^c})((I-\mu_{\Omega^c})^2-\sigma^2_{\Omega^c}) \\ & + k^{(b)}\,\kappa + < \nabla k^{(b)}, \mathbf{n} > . \end{aligned}$$

Mais la vitesse de déplacement V de la courbe Γ n'est autre que $\dfrac{\partial \Gamma}{\partial t}$. Donc Γ est solution de l'équation d'évolution

$$\begin{aligned} \frac{\partial \Gamma}{\partial t} &= -F\,\mathbf{n} \\ &= -\Big(\phi(\sigma^2_\Omega) + \phi'(\sigma^2_\Omega)((I-\mu_\Omega)^2-\sigma^2_\Omega) \\ &\quad -\phi(\sigma^2_{\Omega^c}) - \phi'(\sigma^2_{\Omega^c})((I-\mu_{\Omega^c})^2-\sigma^2_{\Omega^c}) \\ &\quad +k^{(b)}\,\kappa + < \nabla k^{(b)}, \mathbf{n} > \Big)\mathbf{n} \end{aligned} \tag{2.8}$$

Revenons à la convention prise pour définir l'intérieur d'un ensemble de N courbes : compte tenu du fait que Ω_{in} est la réunion d'ouverts disjoints, la méthode peut rencontrer des limitations si plusieurs objets à segmenter ont des propriétés différentes, car les régions risquent de ne pas attirer la courbe avec une force semblable. De manière générale, on peut dire que cette méthode est efficace quand l'ensemble des objets à segmenter est homogène (tous les objets ont le même niveau de gris, la même moyenne, la même variance... selon les descripteurs utilisés).
Dans le cas le plus général possible, on souhaiterait faire évoluer ces N courbes de telle sorte que les intérieurs de chaque courbe forment une partition de D. Pour segmenter K régions $\{\Omega_i\}_{i=1..K}$ formant une partition de D, on a besoin d'imposer que le niveau de gris ait un comportement local suivant une certaine loi de distribution *a priori* (u_0 est constant égal à μ_i sur Ω_i, suit une loi gaussienne de moyenne μ_i et d'écart-type σ_i,...). La difficulté tient à choisir la loi *a priori* la plus adaptée au problème, en l'occurence au type d'image que l'on traite. On fait alors évoluer les N courbes selon N EDP dérivants d'une fonctionnelle généralisant (2.7). Pour plus de détails, on se reportera à [SBFAZ]. Il s'agit en fait d'un problème de *classification* d'images.

2.2.2 Formulation scalaire

Pour formuler une équation d'évolution scalaire à partir de l'équation (2.8), on peut utiliser la méthode des ensembles de niveaux d'Osher-Sethian. On obtient alors

$$\begin{aligned}\frac{\partial u}{\partial t} &= F|\nabla u| \\ &= \Big(\phi(\sigma_\Omega^2) + \phi'(\sigma_\Omega^2)((I-\mu_\Omega)^2 - \sigma_\Omega^2) \\ &\quad -\phi(\sigma_{\Omega^c}^2) - \phi'(\sigma_{\Omega^c}^2)((I-\mu_{\Omega^c})^2 - \sigma_{\Omega^c}^2) \\ &\quad +\mathrm{div}\left(k^{(b)}\frac{\nabla u}{|\nabla u|}\right)\Big)|\nabla u| \qquad (2.9)\end{aligned}$$

Mais même écrite de la sorte, (2.9) n'est pas particulièrement évidente à implémenter, notamment à cause des termes faisant intervenir σ_Ω, σ_{Ω^c}, μ_Ω et μ_{Ω^c}. Et si nous avons réécris l'équation (2.8) sous forme scalaire, nous n'avons pas fait le lien entre ces quantités et u. C'est ce que nous allons faire maintenant, en commençant par réécrire l'énergie J en fonction de u. La modélisation adoptée est reprise des travaux de Laure Blanc-Féraud *et al.* ([SBFAZ]) et également de ceux de T. F. Chan et L. A. Vese ([CV1], [CV2]).

2.2.2.1 Réécriture de l'énergie

On considère que les fonctions $k^{(in)}(\mathbf{x},\Omega)$, $k^{(out)}(\mathbf{x},\Omega^c)$ et $k^{(b)}(\mathbf{x},\Gamma)$ peuvent se prolonger sur tout le domaine D. Soient les fonctions $k_u^{(in)}(\mathbf{x},u) \triangleq k^{(in)}(\mathbf{x},\Omega)$, $k_u^{(out)}(\mathbf{x},u) \triangleq k^{(out)}(\mathbf{x},\Omega^c)$ et $k_u^{(b)}(\mathbf{x},u) \triangleq k^{(b)}(\mathbf{x},\Gamma)$, la distribution de Heaviside H et sa dérivée au sens des distributions δ (la distribution de Dirac), on obtient alors que

$$\begin{aligned}\int_{\Omega_{in}} k^{(in)}(\mathbf{x},\Omega)\,d\mathbf{x} &= \int_\Omega k_u^{(in)}(\mathbf{x},u)H(u)\,d\mathbf{x} \\ \int_{\Omega_{out}} k^{(out)}(\mathbf{x},\Omega^c)\,d\mathbf{x} &= \int_\Omega k_u^{(out)}(\mathbf{x},u)(1-H(u))\,d\mathbf{x}.\end{aligned}$$

Formulé tel quel, en rajoutant l'hypothèse que $k_u^{(in)}$, $k_u^{(out)}$ et $k_u^{(b)}$ sont des fonctions C^∞ à support compact, J est l'action d'une distribution et n'est donc pas Gâteaux-différentiable. On introduit donc des approximations C^1 des distributions de Dirac et de Heaviside.

$$\delta_\alpha(s) = \begin{cases} 0 & \text{si } |s| > \alpha \\ \frac{1}{2\alpha}(1+\cos(\frac{\pi s}{\alpha})) & \text{si } |s| \le \alpha \end{cases}$$

$$H_\alpha(s) = \begin{cases} 1 & \text{si} \quad s > \alpha \\ 0 & \text{si} \quad s < -\alpha \\ \frac{1}{2\alpha}(1+\frac{s}{\alpha}+\frac{1}{\pi}\sin(\frac{\pi s}{\alpha})) & \text{si} \quad |s| \le \alpha. \end{cases}$$

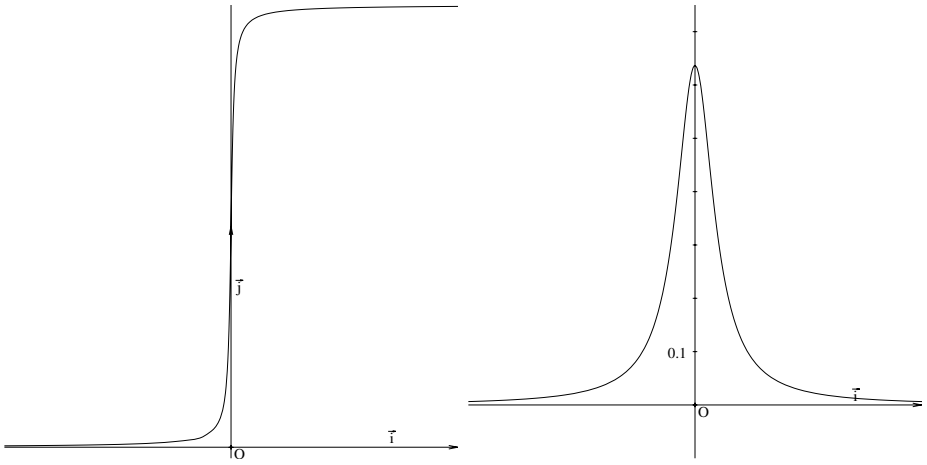


FIG. 2.1: Représentation des approximations C^∞ des distributions de Heaviside et de Dirac $H_\alpha(s) = 1 + \frac{2}{\pi}\arctan(\frac{s}{\alpha})$ et de $\delta_\alpha(s) = \frac{2}{\pi}\frac{\alpha}{\alpha^2+s^2}$.

On peut aussi prendre les approximations C^∞ $H_\alpha(s) = \frac{1}{2}\left(1 + \frac{2}{\pi}\arctan(\frac{s}{\alpha})\right)$ et $\delta_\alpha(s) = \frac{1}{\pi}\frac{\alpha}{\alpha^2+s^2}$. et on remplace les deux distributions par ces approximations (*cf.* figure 2.1). Ainsi, l'énergie devient fonction de u

$$\begin{aligned} J_\alpha(u) = &\int_D k_u^{(in)}(\mathbf{x},u)H_\alpha(u)\,d\mathbf{x} + \int_D k_u^{(out)}(\mathbf{x},u)(1-H_\alpha(u))\,d\mathbf{x} \\ &+ \int_D k_u^{(b)}(\mathbf{x},u)\delta_\alpha(u)|\nabla u|\,d\mathbf{x} \end{aligned}$$

De la même façon, on peut exprimer moyennes et variances approchées en fonction de u :

$$\begin{aligned} \mu_{in,\alpha}(u) &= \frac{\int_D I(\mathbf{x})H_\alpha(u)\,d\mathbf{x}}{\int_D H_\alpha(u)\,d\mathbf{x}} \\ \mu_{out,\alpha}(u) &= \frac{\int_D I(\mathbf{x})(1-H_\alpha(u))\,d\mathbf{x}}{\int_D (1-H_\alpha(u))\,d\mathbf{x}} \\ \sigma^2_{in,\alpha}(u) &= \frac{\int_D (I(\mathbf{x})-\mu_{in,\alpha}(u))^2(1-H_\alpha(u))\,d\mathbf{x}}{\int_D H_\alpha(u)\,d\mathbf{x}} \\ \sigma^2_{out,\alpha}(u) &= \frac{\int_D (I(\mathbf{x})-\mu_{out,\alpha}(u))^2(1-H_\alpha(u))\,d\mathbf{x}}{\int_D (1-H_\alpha(u))\,d\mathbf{x}} \end{aligned}$$

On peut donc ainsi remplacer μ_{Ω,Ω^c} et σ_{Ω,Ω^c} dans (2.9) par leurs nouvelles expressions approchées $\mu_{(in,out),\alpha}$ et $\sigma_{(in,out),\alpha}$. On peut aussi minimiser $J_\alpha(u)$ par

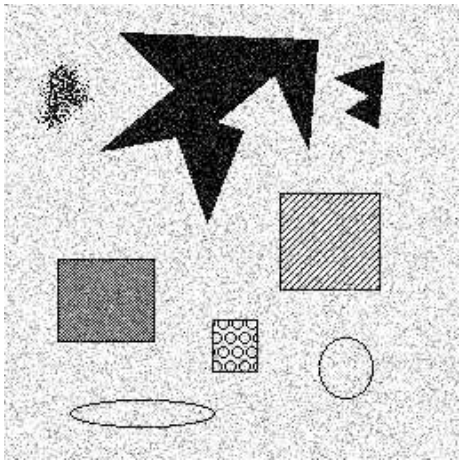


FIG. 2.2: L'image à segmenter.

rapport à u, ce qui va nous donner une EDP différente de (2.9)

$$\begin{aligned}\frac{\partial u}{\partial t} &= F|\nabla u| \\ &= \Big(\phi(\sigma^2_{in,\alpha}(u)) + \phi'(\sigma^2_{in,\alpha}(u))((I-\mu_{in,\alpha}(u))^2 - \sigma^2_{in,\alpha}(u)) \\ &\quad -\phi(\sigma^2_{out,\alpha}(u)) - \phi'(\sigma^2_{out,\alpha}(u))((I-\mu_{out,\alpha}(u))^2 - \sigma^2_{out,\alpha}(u)) \\ &\quad +\mathrm{div}\left(k^{(b)}\frac{\nabla u}{|\nabla u|}\right)\Big)\delta_\alpha(u). \qquad (2.10)\end{aligned}$$

La différence entre les deux approches est philosophique et réside dans le choix de l'extension à tout le domaine d'une fonction définie sur la courbe Γ. Barlaud *et al.* avaient choisi d'implémenter (2.9) avec un schéma explicite et la réinitialisation de la fonction distance, alors que nous avons choisi de plutôt implémenter (2.10) avec un schéma semi-implicite (*cf.* section 2.4.2), ce qui permet d'éviter en pratique la réinitialisation de la fonction distance, mais au prix de calculs plus coûteux.

2.3 Résultats et comparaison des méthodes

On va présenter ici quelques résultats montrant les différences entre des snakes basés régions comme ceux de Barlaud *et al.* ou ceux de Chan et Vese et des snakes purement orientés contours comme ceux de Caselles, Kimmel et Sapiro.

Comme nous l'avons dit plus haut, si on ne rajoute pas de terme de minimisation d'aire dans le modèle des contours actifs géodésiques, la courbe a tendance à

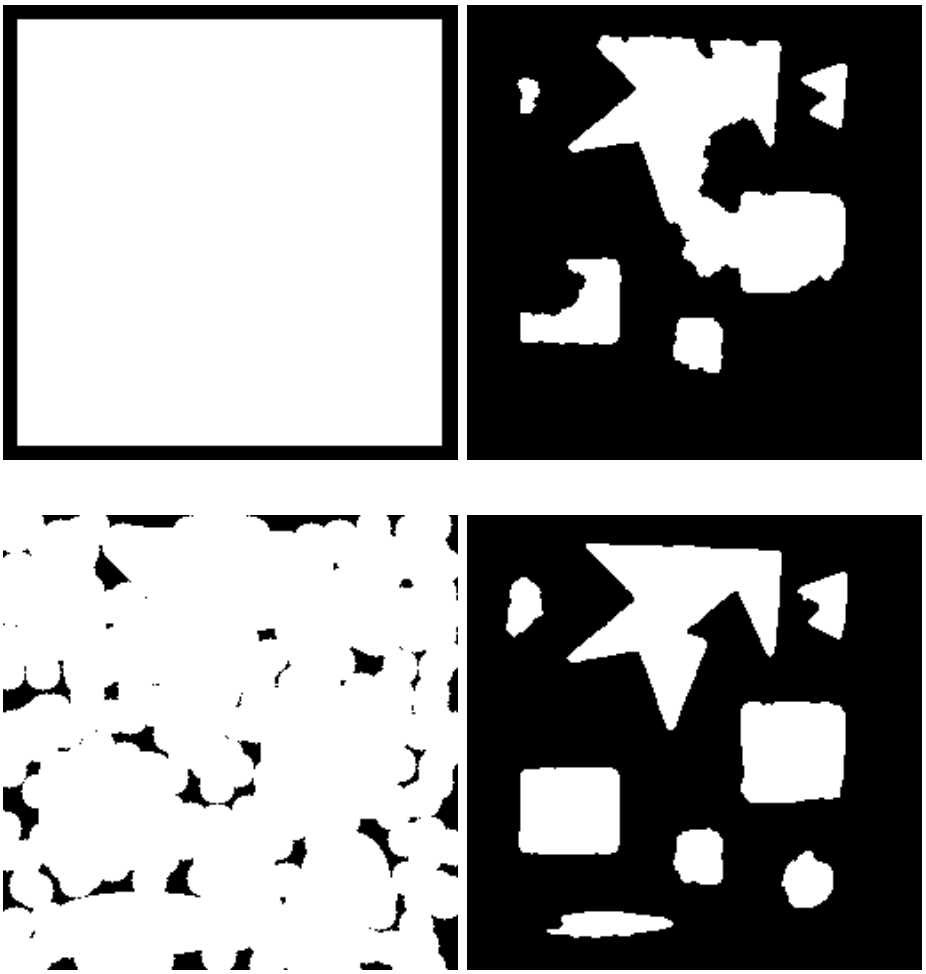


FIG. 2.3: Les masques d'initialisation (images de gauche) et de résultats obtenus par les contours actifs géodésiques sur l'image précédente. Le premier résultat est obtenu pour 25000 itérations, $\delta t = 2$, $\nu = 0,25$, l'image est convoluée par un noyau gaussien G_σ avec $\sigma = 0,8$. Le contour perd l'attache aux bords au bout d'un certain nombre d'itérations. Le deuxième résultat (image en bas à droite) est obtenu à partir d'une initialisation plus proche, nécessite moins d'itérations (10000) et l'attache aux bords est meilleure.

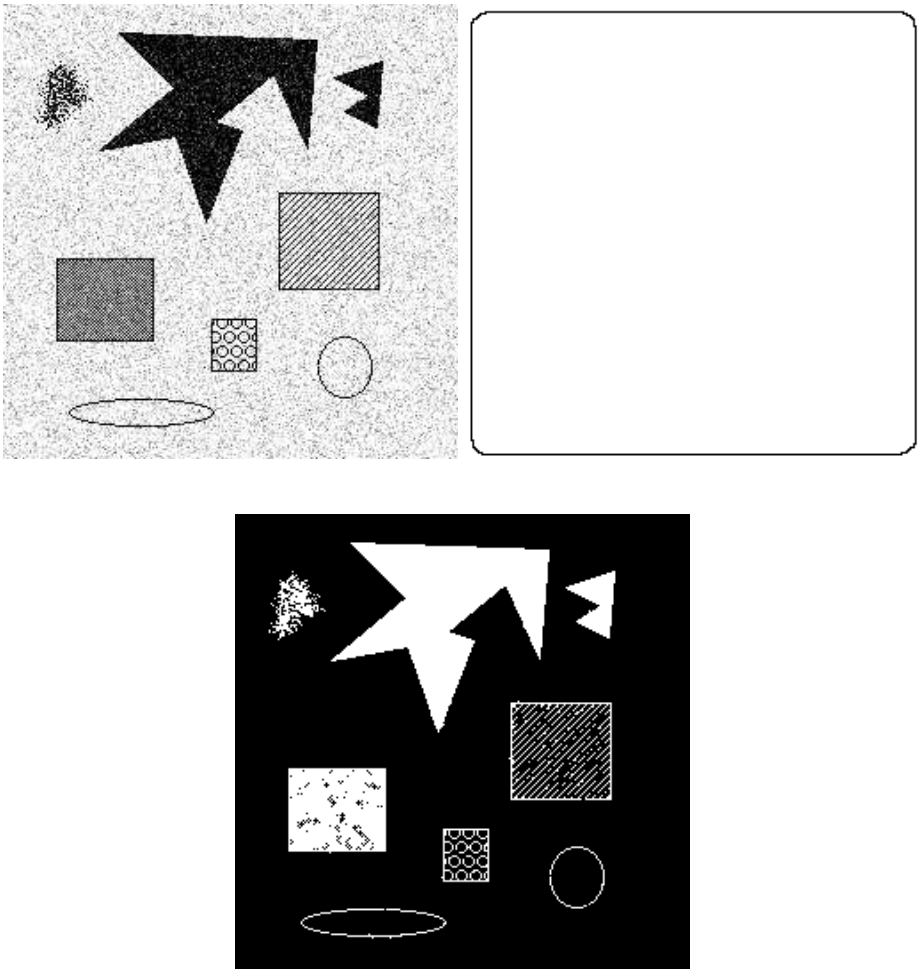


FIG. 2.4: Résultat de l'algorithme de contours actifs de Chan et Vese. En haut : l'image et le contour initial (en noir), l'ensemble de niveau 0 de u (paramètres : 1500 itérations, $\delta t = 0,05$). La courbe arrive à se scinder correctement pour segmenter l'ensemble des objets, y compris l'intérieur des ensembles texturés. On remarque toutefois les effets du bruit sur les zones texturées.

converger vers l'enveloppe convexe des objets et de plus son évolution est assez lente. Cela nous donne donc, grâce à l'écriture en *level sets* de l'équation (2.1)

$$\frac{\partial u}{\partial t} = |\nabla u| \mathrm{div}\left(g\frac{\nabla u}{|\nabla u|}\right) + \nu g|\nabla u|.$$

L'ajout du terme $\nu g|\nabla u|$ a pour effet d'accélérer l'évolution de u tout en évitant de l'accélérer trop lorsque la courbe s'approche d'un bord, la présence de g atténuant en effet ce terme près des bords. On présente **en figure 2.3** deux résultats obtenus respectivement pour les contours actifs géodésiques pour deux initialisations : la première est une initialisation très large, la deuxième est obtenue en seuillant l'image et en dilatant le résultat. Il s'agit donc d'une initialisation plus proche du résultat. On voit que le meilleur résultat est obtenu pour l'initialisation la plus proche du résultat. En effet, l'équation des géodésiques tend à réduire la courbe jusqu'à disparition : au bout d'un certain nombre d'itérations, il y a une perte d'attache aux bords de l'image. C'est ce qui est observé en **figure 2.3**. Toutefois une initialisation plus proche permet de segmenter correctement les objets si le nombre d'itérations est judicieusement choisi.
Le résultat présenté sur la **figure 2.4** illustre bien les propriétés des modèles de contours actifs basés régions. Ici, nous avons utilisé le modèle de Chan et Vese, ce qui revient à prendre $\phi(s) = s$ dans le modèle de Barlaud *et al.* L'image à segmenter est presque un cas idéal d'image donnant un bon résultat pour cet algorithme : bien que l'ensemble des objets à segmenter soit non connexe, il est de niveau de gris uniforme (noir) et assez contrasté par rapport au fond. On voit aussi que le contour rentre à l'intérieur des objets et notamment que l'algorithme sépare correctement les hachures du carré.
Si le niveau de gris présentaient plus de deux modes caractéristiques, l'utilisation de cet algorithme tel quel serait discutable, car on voudrait alors autant de régions que de modes. Le mieux serait alors d'utiliser la version de l'algorithme généralisé à une segmentation en plusieurs régions.
Le contour arrive bien à se scinder, son intérieur étant alors défini comme la réunion des intérieurs de ses composantes connexes. Une autre propriété assez intéressante de cet algorithme est de faire converger le contour sans le faire s'effondrer comme peuvent le faire les algorithmes de contours actifs géodésiques ou géométriques, car si on augmente le nombre d'itérations, on constate que le contour a bien atteint son état d'équilibre et n'a pas bougé.

Comme on peut le voir **en figure 2.3**, les résultats entre les contours actifs géométriques et géodésiques sont assez semblables, la seule petite différence que l'on peut constater est une certaine régularisation des pointes de certains objets sur les résultats relatifs au modèle géométrique. On l'a déjà dit, les parties texturées sont mal segmentées car la courbe est bloquée par le premier contour rencontré. De plus, en ce qui concerne le carré hachuré, les hachures sont tellement denses que la convolution de l'image par un noyau gaussien induit un

gradient fort partout dans cette zone, ainsi, il est quasiment impossible de la segmenter correctement en séparant les hachures du reste.

2.4 Revue des implémentations

Il existe plusieurs façons d'implémenter une EDP dépendant de la courbure. Nous allons principalement présenter deux approches. La première est issue des travaux d'Alvarez ([Al]) et de Cohignac, Eve, Guichard, Lopez et Morel ([CEGLM]). Elle provient de la discrétisation de l'opérateur $|\nabla u|\text{div}\left(\frac{\nabla u}{|\nabla u|}\right)$. La deuxième vient d'une discrétisation de l'opérateur de diffusion $\text{div}\,(d\nabla u)$ proposée par Malik et Perona dans [PM].

2.4.1 A partir du schéma d'Alvarez et Guichard

Soit ξ un vecteur normé tel que $\nabla u \cdot \xi = 0$. ξ est donc le vecteur normal à ∇u si ce vecteur est non nul. On peut montrer que l'on a $|\nabla u|\text{div}\left(\frac{\nabla u}{|\nabla u|}\right) = u_{\xi\xi}$. Afin d'avoir la consistance du schéma numérique, on cherche les coefficients λ_0, λ_1, λ_2, λ_3, λ_4 tels que ($u_{i,j} \triangleq u(ih, jh)$ où h désigne le pas d'espace)

$$u_{\xi\xi} = \frac{1}{h^2}(-4\lambda_0 u_{i,j} + \lambda_2(u_{i+1,j} + u_{i-1,j}) + \lambda_3(u_{i-1,j-1} + u_{i+1,j+1}) + \lambda_4(u_{i-1,j+1} + u_{i+1,j-1})) + o(1)$$

On peut alors en déduire (*cf.* [CEGLM]), en introduisant θ tel que $\xi = (\cos(\theta), \sin(\theta))$,

$$\begin{aligned}
\lambda_1 &= 2\lambda_0 - \sin^2\theta \\
\lambda_2 &= 2\lambda_0 - \cos^2\theta \\
\lambda_3 &= -\lambda_0 + 0.5(1 + \sin\theta\cos\theta) \\
\lambda_3 &= -\lambda_0 + 0.5(1 - \sin\theta\cos\theta).
\end{aligned}$$

Et dans [CEGLM], les auteurs choisissent $\lambda_0 = 0.5 - \cos^2\theta + \cos^4\theta$.
Dans le cadre des contours actifs géométriques, les auteurs de [CCCD] suivent l'idée introduite par Alvarez dans [Al] : écrire les termes en u de l'expression de $u_{\xi\xi}$ au temps $n+1$ et les termes en λ au temps n, ce qui donne donc

$$\begin{aligned}
\frac{u^{n+1} - u^n}{\delta t} =& \nu g|\nabla u^n| + g[\frac{1}{h^2}(-4\lambda_0^n u_{i,j}^{n+1} + \lambda_2^n(u_{i+1,j}^{n+1} + u_{i-1,j}^{n+1}) \\
&+ \lambda_3^n(u_{i-1,j-1}^{n+1} + u_{i+1,j+1}^{n+1}) + \lambda_4^n(u_{i-1,j+1}^{n+1} + u_{i+1,j-1}^{n+1}))].
\end{aligned}$$

Il s'agit donc d'un schéma semi-implicite (les poids sont au temps n et les valeurs de u au temps $n+1$) de l'équation de la courbure moyenne. Les schémas implicites

ou semi-implicites sont connus pour donner une meilleure stabilité au schéma numérique que des schémas explicites qui limitent la valeur du pas de temps. Il en coûte toutefois de plus longs calculs (résolution de systèmes linéaires).
Le système linéaire en u^{n+1} peut alors être résolu par une méthode itérative, mais en pratique on peut approcher la solution exacte en faisant une itération de la méthode de Jacobi en parcourant l'image d'en haut à gauche jusqu'en bas à droite, puis une itération en sens inverse (*cf.* [Al]).

2.4.2 A partir du schéma de Malik et Perona

Comme dans [WK], nous allons écrire une EDP générique sous la forme

$$\frac{\partial u}{\partial t} = a(x)|\nabla u|\mathrm{div}\left(b(x)\frac{\nabla u}{|\nabla u|}\right) + |\nabla u|\nu g(x).$$

Cette formulation a l'avantage de pouvoir contenir les différentes EDP remarquables que nous avons citées : si $a = 1$ et $b = g$, on a les contours actifs géodésiques ; si $a = g$ et $b = 1$, on a les contours actifs géométriques ; et si $a = 1$, $b = 1$ et $k = 0$, on a l'équation de la courbure moyenne.
L'idée suivie par les auteurs de [WK] est d'utiliser la discrétisation de l'opérateur $\mathrm{div}\,(d\nabla u)$ proposée par Malik et Perona dans [PM]. Partant de

$$\mathrm{div}\,(d\nabla u) = \frac{\partial}{\partial x}\left(d\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(d\frac{\partial u}{\partial y}\right),$$

on approche $\dfrac{\partial f}{\partial x}$ par $\frac{f_{i+\frac{1}{2},j}-f_{i-\frac{1}{2},j}}{h}$ et $\dfrac{\partial f}{\partial y}$ par $\frac{f_{i,j+\frac{1}{2}}-f_{i,j-\frac{1}{2}}}{h}$, h étant le pas de discrétisation spatial. On trouve alors

$$\begin{aligned}(\mathrm{div}\,(d\nabla u)) =& \frac{1}{h^2}[d_{i+\frac{1}{2},j}u_{i+1,j} + d_{i-\frac{1}{2},j}u_{i-1,j} + d_{i,j+\frac{1}{2}}u_{i,j+1} + d_{i,j-\frac{1}{2}}u_{i,j-1}\\ &- (d_{i+\frac{1}{2},j} + d_{i-\frac{1}{2},j} + d_{i,j+\frac{1}{2}} + d_{i,j-\frac{1}{2}})] + o(1).\end{aligned}$$

On peut aussi faire intervenir les éléments diagonaux et pondérer l'expression obtenue ci-dessus avec l'expression incluant les éléments diagonaux. Pour de plus amples détails sur ce point, on se reportera à [AK].
Pour estimer les $d_{i\pm\frac{1}{2},j}$ (*resp.* $d_{i,j\pm\frac{1}{2}}$), on peut prendre la moyenne $\frac{d_{i,j}+d_{i\pm1,j}}{2}$ (*resp.* $\frac{d_{i,j}+d_{i,j\pm1}}{2}$) ou la moyenne harmonique $\frac{2}{\frac{1}{d_{i,j}}+\frac{1}{d_{i\pm1,j}}}$ (*resp.* $\frac{2}{\frac{1}{d_{i,j}}+\frac{1}{d_{i,j\pm1}}}$). Les deux sont justifiées par un développement de Taylor du premier ordre.
Pour l'EDP écrite ci-dessus, on a $d = \frac{b}{|\nabla u|}$, et il devient intéressant de considérer la moyenne harmonique. En effet, la moyenne arithmétique est donnée par

$$\frac{\left(\frac{b}{|\nabla u|}\right)_{i',j'} + \left(\frac{b}{|\nabla u|}\right)_{i,j}}{2}$$

ce qui pose un problème en cas d'annulation de ∇u.
En revanche, la moyenne harmonique

$$\frac{2}{\left(\frac{|\nabla u|}{b}\right)_{i',j'} + \left(\frac{|\nabla u|}{b}\right)_{i,j}}$$

est définie dès que $(\nabla u)_{i,j} \neq (0,0)^T$.

Le terme $|\nabla u| kg(x)$ est discrétisé selon la méthode du min-mod (*cf.* [WK]), ce qui donne

$$|\nabla^+ u|^n_{i,j} = (min(\delta^{-x} u^n_{i,j}, 0)^2 + max(\delta^{+x} u^n_{i,j}, 0)^2 + min(\delta^{-y} u^n_{i,j}, 0)^2 + max(\delta^{+y} u^n_{i,j}, 0)^2)^{\frac{1}{2}} \text{ si } k > 0,$$

les termes en d sont écris au temps n et les termes en u au temps $n+1$, ce qui permet d'écrire

$$(Id - \delta t M)[u^{n+1}_{i,j}] = [u^n_{i,j}] + \nu g(x)[\nabla^+ u^n_{i,j}] \tag{2.11}$$

où A est la matrice

$$M = (M_{i',j'}) = \begin{cases} -a_{i,j}|\nabla u^n_{i,j}| \sum\limits_{(i',j') \in \mathcal{N}(i,j)} \frac{2}{\left(\frac{|\nabla u|}{b}\right)_{i',j'} + \left(\frac{|\nabla u|}{b}\right)_{i,j}} & \text{si } (i',j') = (i,j) \\ a_{i,j}|\nabla u^n_{i,j}| \frac{2}{\left(\frac{|\nabla u|}{b}\right)_{i',j'} + \left(\frac{|\nabla u|}{b}\right)_{i,j}} & \text{si } (i',j') \in \mathcal{N}(i,j); \end{cases}$$

$\mathcal{N}(i,j)$ est le voisinage du pixel (i,j) constitué des quatre plus proches voisins et $[u^{n+1}_{i,j}]$ est un vecteur colonne obtenu par réorganisation des $u^{n+1}_{i,j}$. Plutôt que de résoudre (2.11), on résout indépendamment un système suivant l'axe x et un système suivant l'axe y : on a donc deux matrices M_x et M_y définies à partir du voisinage des deux plus proches pixels suivant chaque axe. Ce procédé est appelé l'*Additive Operator Splitting* (AOS). L'avantage d'une telle séparation des traitements selon chaque axe est de manipuler des matrices tridiagonales. La solution du système est alors

$$[u^{n+1}_{i,j}] = \frac{1}{2} \sum_{l \in \{x,y\}} (Id - \delta t M_l)^{-1} ([u^n_{i,j}] + \nu g(x)[\nabla^+ u^n_{i,j}])$$

et l'inversion est réalisée soit par une méthode itérative, soit par l'algorithme de Thomas calculant l'inverse d'une matrice tridiagonale. On peut montrer aisément que pour $k = 0$ le schéma est inconditionnellement stable et qu'il vérifie le principe du maximum (les matrices $Id - \delta t M_l$ sont à terme positif et la somme de chaque ligne donne 1). Certes la fonction u ne reste pas une fonction distance au cours de son évolution mais la propriété de stabilité énoncée ci-dessus est très intéressante numériquement ; on peut aussi combiner le schéma avec les techniques classiques de réinitialisation.

2.4.3 A partir d'opérateurs morphologiques

La théorie du comportement asymptotique des opérateurs morphologiques classiquement utilisés en traitement d'images nous dit que l'itération d'un filtre médian tend à approcher l'évolution suivant la courbure moyenne, et que l'opérateur de dilatation tend à approcher l'évolution suivant l'équation $\frac{\partial u}{\partial t} = |\nabla u|$.

Théorème 2 (Guichard-Morel [GM]) *Soit u une fonction $C^3(\mathbb{R}^2, \mathbb{R})$. Soit μ_h la mesure uniforme sur le disque $D(0,h)$*

$$\mu_h = \frac{\mathbf{1}_{D(0,h)}}{\pi h^2}$$

On définit l'opérateur médian par

$$\mathrm{med}_{D(0,h)} u(x,y) = \inf_{\mu_h(B) \geq \frac{1}{2},\ B \subset D(0,h)} \sup_{(x',y') \in B} u(x+x', y+y')$$

Alors

$$\mathrm{med}_{D(0,h)} u(x,y) = u(x,y) + \frac{1}{6} |\nabla u| \mathrm{curv}(u) h^2 + O_{x,y}(h^3).$$

On peut donc en déduire le théorème de convergence suivant

Théorème 3 (Guichard-Morel [GM]) *Soit $T_h^n u_0 = \mathrm{med}^n_{D(0,h)} u_0$. On a $T_h^n u_0(\mathbf{x}) \to u(\mathbf{x}, t)$ lorsque $h \to 0$, $n \to \infty$ et $\frac{nh^2}{6} \to t$ avec u qui vérifie l'équation de la courbure moyenne $\frac{\partial u}{\partial t} = |\nabla u| \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)$.*

Un autre résultat dû à F. Catté, F. Dibos et G. Koepfler ([CDK]) nous dit que la moyenne d'un inf-sup et d'un sup-inf tend aussi à approcher un mouvement par courbure moyenne.

Théorème 4 *Soit $\mathcal{B}$ l'ensemble des segments du plan de longueur 2 centrés en 0. On définit*

$$SI_h u(\mathbf{x}) = \sup_{B \in \mathcal{B}} \inf_{\mathbf{y} \in \mathbf{x} + hB} u(\mathbf{y})$$

et

$$IS_h u(\mathbf{x}) = \inf_{B \in \mathcal{B}} \sup_{\mathbf{y} \in \mathbf{x} + hB} u(\mathbf{y}).$$

Pour une fonction $u \in \mathcal{C}^3(\mathbb{R}^2)$, on a le résultat

$$\frac{1}{2}\left(IS_h + SI_h\right) u(\mathbf{x}_0) = u(\mathbf{x}_0) + h^2 \frac{1}{4} |\nabla u| \mathrm{curv}(u)(\mathbf{x}_0) + O(h^3)$$

si $|Du|(\mathbf{x}_0) \neq 0$.

De là un théorème de convergence similaire au précédent.

Théorème 5 (Catté, Dibos, Koepfler [CDK]) *On conserve les définitions de SI_h et IS_h du théorème précédent.*
Soit $T_h^n u_0 = \left(\frac{SI_h+IS_h}{2}\right)^n u_0$. On a $T_h^n u_0(\mathbf{x}) \to u(\mathbf{x},t)$ lorsque $h \to 0$, $n \to \infty$ et $\frac{nh^2}{4} \to t$ avec u qui vérifie l'équation de la courbure moyenne $\dfrac{\partial u}{\partial t} = |\nabla u|\mathrm{div}\left(\dfrac{\nabla u}{|\nabla u|}\right)$.

Théorème 6 *Soit la dilatation D_t (*resp. *l'érosion E_t) d'élément structurant le disque unité $D(0,1)$*

$$D_t u_0(x) = \sup_{(x',y')\in D(0,1)} u_0(x - tx', y - ty')$$

$$E_t u_0(x) = \inf_{(x',y')\in D(0,1)} u_0(x + tx', y + ty').$$

*Alors si on note $u(t,x) = D_t u_0(x)$ (*resp. *$u(t,x) = E_t u_0(x)$), $u(t,x)$ satisfait à*

$$\frac{\partial u}{\partial t} = +|\nabla u|$$

*(*resp. $\dfrac{\partial u}{\partial t} = -|\nabla u|$*).*

Partant de ces approximations des EDP par l'itération d'opérateurs morphologiques, Guichard et Morel ([GM]) écrivent un schéma non morphologique, mais qui satisfait le principe du maximum pourvu que $\delta t g(x,y) \leq 1$ et $\delta t|\nabla g|(x,y) \leq 1$. La discrétisation de

$$\frac{\partial u}{\partial t} = g|\nabla u|\left(\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) + 1\right) + \nabla g \cdot \nabla u$$

est donnée par

$$u^{n+1} = u^n + \delta t g(med_{\sqrt{6}}(u^n) - u^n + D_1(u^n) - u^n) + \delta t\left(D_1(u^n)(x - \frac{\partial g}{\partial x}, y - \frac{\partial g}{\partial y}) - D_1(u^n)(x,y)\right).$$

2.4.4 Implémentation des EDPs avec fonctions de Heaviside

Comme nous l'avons vu, pour certains problèmes de minimisation (Chan–Vese, Aubert–Barlaud–Jehan-Besson,...), on est conduit à étudier des EDP de la forme

$$\frac{\partial u}{\partial t} = \delta_\epsilon(u(x))\left(f(x) + \mu\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)\right)$$

où f est une fonction dépendant du point.
En fait, on peut reprendre la technique que nous avons présenté dans la section sur la discrétisation de Malik et Perona. En effet, $\delta_\epsilon(u)$ est strictement positif et

inférieur à $\frac{1}{\epsilon}$. On utilise alors les schémas AOS sur l'opérateur divergence. Bien qu'à notre connaissance, cela n'ait été implémenté de cette façon par aucun auteur, les résultats se sont montrés aussi corrects que les résultats montrés dans [CV1] (voir section 2.3) et [ABJB2] (voir section 3.2.1.1). De plus, cette implémentation permet d'éviter la réinitialisation de la fonction distance qui est bien souvent nécessaire lorsque l'on utilise un schéma explicite. Dans nos expériences nous nous sommes contentés d'extraire les lignes de niveau 0 de la fonction u après les calculs du schéma itératif.

Chapitre 3

Segmentation d'objets en mouvement à caméra fixe

Dans le cadre de la segmentation d'objets en mouvement dans une séquence vidéo filmée à caméra fixe, nous allons présenter différents modèles ainsi que leur mise en oeuvre algorithmique.
La première section de ce chapitre est consacrée à une méthode de segmentation conjointement à un calcul du fond statique. Cette méthode a été introduite par Aubert, Deriche et Kornprobst. Nous en décrivons les avantages et les inconvénients, expériences à l'appui.
La deuxième partie de ce chapitre a pour but de proposer des algorithmes plus simples et plus rapides pour la minimisation de certaines fonctionnelles initialement écrites dans un cadre continu. Pour cela, nous faisons le lien entre énergies discrètes et continues et appliquons l'algorithme des modes conditionnels itérés ou ICM (Iterated Conditional Modes). Nous présentons des applications sur un modèle récent de segmentation vidéo par comparaison entre l'image et le fond et sur un modèle de suivi d'objets.
Dans la dernière partie de ce chapitre, nous nous intéressons à l'utilisation du flot optique pour la détection d'objets en mouvement. Nous présentons d'abord les problème liés à la détermination du flot optique, ensuite les quelques modèles de segmentation à partir de la connaissance du flot optique qui ont été proposés ; enfin, nous présentons un nouveau modèle qui s'apparente à ceux utilisés pour comparer une image avec un fond donné.

3.1 Segmentation conjointe fond/objets en mouvement

Dans [ADK2], Aubert, Deriche et Kornprobst proposent de calculer le fond statique ainsi que les objets en mouvement dans une sequence d'images en minimisant la fonctionnelle (on reprend les notations de [ADK2]). L'idée de segmenter

simultanément le mouvement et d'extraire le fond vient de l'interdépendance entre les deux : si l'on réussit à extraire le fond, c'est que d'une certaine manière, même si ce n'est pas forcément explicite, on a segmenté le mouvement, et inversement on a besoin du fond pour segmenter le mouvement. A partir de là, l'idée de Kornprobst est de chercher le fond B et des images C_h à valeurs dans $[0,1]$ qui réalisent le minimum de la fonctionnelle suivante

$$E(B, C_1, ..., , C_T) = \sum_{h=1}^{T} \int_D C_h^2 (B - I_h)^2 \, dx + \alpha_c \sum_{h=1}^{T} \int_D (C_h - 1)^2 \, dx + \alpha_b^r \int_D \phi_1(|\nabla B|) \, dx + \alpha_c^r \sum_{h=1}^{T} \int_D \phi_2(|\nabla C_h|) \, dx$$

On choisit pour ϕ_1 et ϕ_2 la fonction d'hypersurface minimale

$$\phi(t) = \sqrt{1 + t^2} - 1.$$

$(I_h)_{h=1}^{h=T}$ sont les T images de la séquence, B le fond à estimer et $(C_h)_{h=1}^{h=T}$ caractérise l'écart entre $(I_h)_{h=1}^{h=T}$ et B. Idéalement, x appartient à une zone en mouvement quand $C_h(x) = 0$, ou bien x appartient au fond lorsque $C_h(x) = 1$. Afin de décider entre ces deux cas, on se fixe un seuil s : si $C_h < s$, $C_h = 0$, $C_h = 1$ sinon. Les auteurs démontrent un résultat d'existence et d'unicité pour le minimum de cette fonctionnelle, on se reportera à [ADK2] pour le détail de cette démonstration.
Cherchons à interpréter qualitativement la fonctionnelle E.

- Le premier terme impose aux pixels tels que la variation $(B - I_h)^2$ entre les images et le fond soit grande d'appartenir aux zones mobiles, puisque pour minimiser le premier terme il faut que C_h soit faible, donc proche de zéro.
- Les deux derniers termes sont des termes de régularisation qui réalisent un lissage anisotropique du fond et des images C_h de façon à lisser les zones homogènes tout en préservant les discontinuités.

L'algorithme de minimisation de la fonctionnelle décrit dans [ADK2] est également sensible aux paramètres α_c, α_b^r et α_c^r : une valeur de α_c trop faible va entraîner une minimisation du premier terme au détriment du deuxième, et donc forçer C_h à zéro, et ce même si l'écart entre B et I_h n'est pas dû au mouvement. A l'inverse, si α_c est trop grand, on risque une mauvaise détection du mouvement, c'est-à-dire que C_h est forcé à 1 (on considère qu'il n'y a pas mouvement) dans des zones de mouvement afin d'annuler le deuxième terme.
Cette méthode de segmentation et de recherche du fond est particulièrement bien adaptée pour une séquence où les objets présentent des déplacements importants, de sorte que le fond soit visible pendant une grande partie de la durée de la séquence. En clair, si un pixel du fond est occulté pendant plus de 60% de la durée totale de la séquence, le fond risque d'être mal reconstitué. A cause du terme de régularisation du fond, on voit que l'algorithme fonctionnera mal

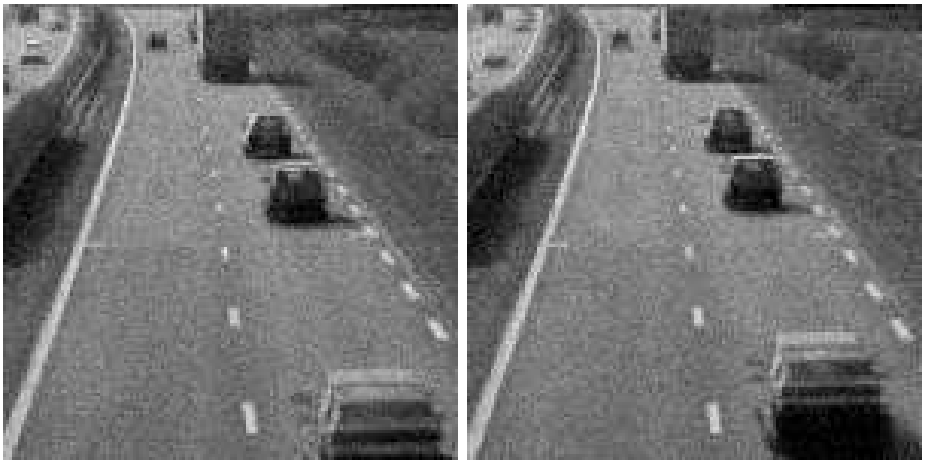
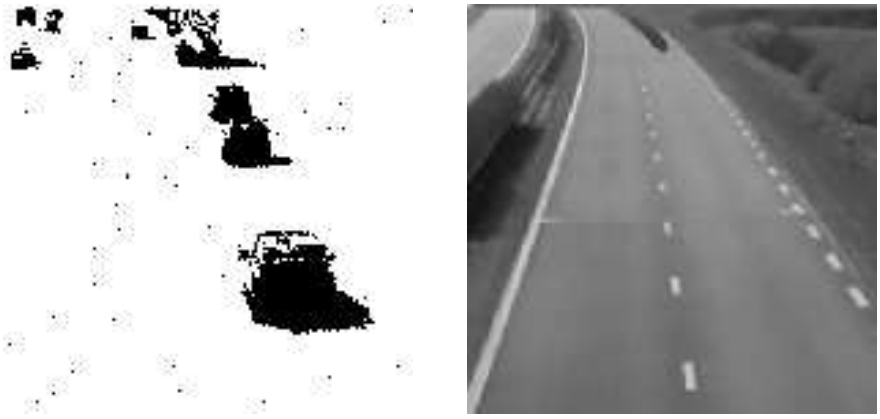


FIG. 3.1: Deux images successives de la séquence d'autoroute.

sur une séquence où le fond est texturé, mais c'est là une limitation courante pour une méthode de segmentation. Sur les **figures 3.2 et 3.4**, on présente les résultats obtenus sur des séquences qui vérifient bien l'hypothèse de bonne visibilité du fond, ainsi les résultats apparaissent comme tout à fait corrects.
Une autre limitation est dûe au fait que l'algorithme décide si un pixel appartient au fond ou bien aux objets en mouvement en fonction de l'écart $(B - I_h)^2$ de niveau de gris entre l'image et le fond. Par conséquent si certains objets en mouvement sont peu contrastés par rapport au fond, il va être difficile de les détecter.
Les différentes limitations de l'algorithme décrites ci-dessus sont bien illustrées sur la séquence des taxis de Hambourg[1] (***cf.* figure 3.6**). Il s'agit d'un cas typique de séquence présentant les difficultés énoncées précédemment. On présente ci-dessous la première et la dernière image de la séquence, ainsi que l'une des images reconstituée ainsi que le fond. On présente également les résultats obtenus sur la séquence d'autoroute : c'est un exemple de séquence sur laquelle l'algorithme donne de très bon résultats.
Dans certains cas, cette détection des objets en mouvement et du fond présente un avantage sur un algorithme de contour actif utilisant le résultat du calcul de flot optique comme on se propose de le faire (*cf.* section 3.4 page 70) : en effet, un cas pratique d'utilisation de l'algorithme d'Aubert, Deriche et Kornprobst est le traitement d'une séquence avec de grands déplacements ; sur de telles séquences, le calcul du flot optique par des méthodes classiques n'est plus valide, puisque ces méthodes nécessitent l'hypothèse de faibles déplacements. Mémin et Perez dans [MP], ou Alvarez, Weickert et Sanchez dans [AWS], ont

[1] disponible sur le site FTP `ftp://ftp.csd.uwo.ca` sous le répertoire `pub/vision`



(a) Objets en mouvements pour une image extraite de la séquence d'autoroute.

(b) Image du fond.

FIG. 3.2: Résultats de la segmentation par méthode variationnelle d'Aubert, Deriche et Kornprobst sur la séquence d'autoroute ($\alpha_c = 300$, $\alpha_b^r = 40$, $\alpha_c^r = 35$). Cas quasi idéal où les objets en mouvement se déplacent assez vite pour que le fond soit visible pendant la majeure partie du temps. Seule une zone en haut de l'image est trop occultée par les voitures à cause de l'effet de perspective. On voit donc une trainée noire à cet endroit sur l'image du fond.

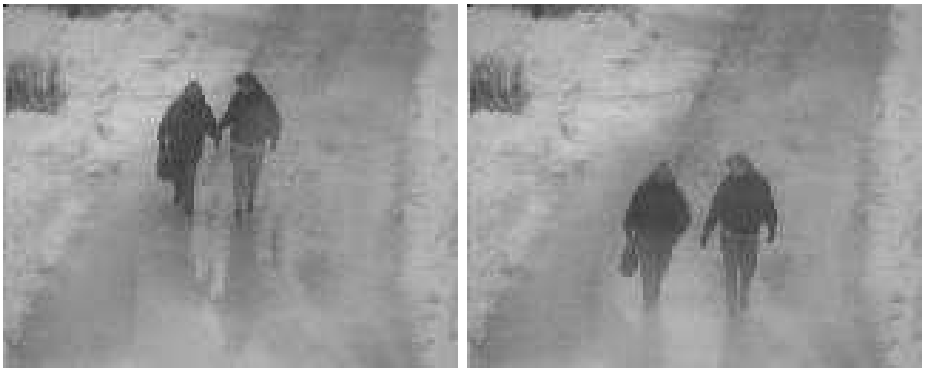
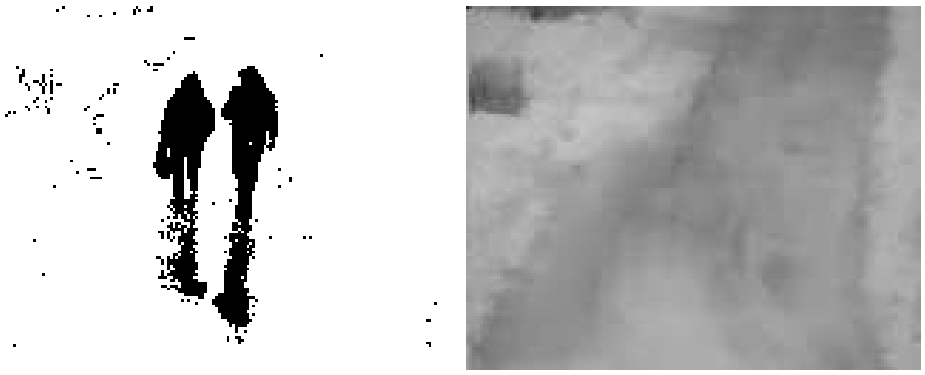


FIG. 3.3: Deux images successives de la séquence "finlandaise".



(a) Objets en mouvements pour la 30^{e} image de la séquence "finlandaise".

(b) Image du fond.

FIG. 3.4: Résultats de la segmentation par méthode variationnelle d'Aubert, Deriche et Kornprobst sur la séquence "finlandaise" ($\alpha_c = 100$, $\alpha_b^r = 10$, $\alpha_c^r = 10$). Cas où l'algorithme s'applique bien : Les personnages se déplacent assez vite pour que le fond soit visible pendant la majeure partie de la durée de la séquence.

FIG. 3.5: Première et dernière image de la séquence des taxis de Hambourg.

(a) Objets en mouvement pour une image extraite de la séquence des taxis.

(b) Image du fond.

FIG. 3.6: Résultats de la segmentation par méthode variationnelle d'Aubert, Deriche et Kornprobst sur la séquence des taxis de Hambourg ($\alpha_c = 300$, $\alpha_b^r = 40$, $\alpha_c^r = 35$). Cas où l'algorithme s'applique mal : les objets se déplacent trop lentement et la durée de la séquence est trop faible pour que le fond puisse être déterminé de manière fiable.

proposé une approche multirésolution permettant de calculer le flot optique dans le cas de déplacements grands devant la taille du masque de discrétisation utilisé pour approcher les dérivées spatiales des images.

3.2 Application de l'ICM et des contours actifs à la segmentation et au tracking d'objets en mouvement

L'approche que nous venons de présenter combine la détermination du fond avec la détermination des objets en mouvement. Les autres approches qui ont été envisagées utilisent davantage des données déterminées par ailleurs, mais on cherche à les améliorer en introduisant des critères de régularité qui sont intégrés à des modèles de contours actifs. Notre but est ici d'utiliser des modèles dans leur version discrète afin de pouvoir leur appliquer l'algorithme des modes conditionnels itérés (ICM).
A notre connaissance, l'équivalence entre certains modèles de contours actifs et des modèles déterministes discrets formulés dans le cadre des champs de Markov a été peu notée, bien que des travaux de Chambolle ([Cham]) donnent une preuve de convergence des solutions du problème discret de Geman et Geman ([Ge]) vers les solutions d'une version anisotropique du problème de Mumford et Shah ([MS1]). Récemment, Chan et Song dans [SC] appliquèrent un algorithme

discret qui n'est autre que l'ICM (l'équivalence de leur algorithme avec l'ICM n'a toutefois pas été notée) à la fonctionnelle de Mumford et Shah constant par morceaux.
Ici, nous reprenons cette idée d'appliquer l'ICM pour d'autres types de fonctionnelle et nous la justifions par un rappel sur les champs de Markov. La section 3.2.1 est consacrée à un modèle de segmentation des objets en mouvement à partir du fond et la section 3.2.2 à un modèle de suivi d'objets ou *tracking*.

3.2.1 Comparaison des images avec un fond donné

3.2.1.1 Formulation continue, contours actifs basés régions, EDPs

Cette fois, on suppose le fond statique connu. Il a pu être estimé par l'algorithme d'Aubert, Deriche et Kornprobst pour segmenter les objets en mouvement ou alors par un autre type de calcul (filtre médian temporel par exemple) ou bien tout simplement par observation de la scène à un moment où «il ne se passe rien». Dans [ABJB2], Aubert, Barlaud et Jehan-Besson ont proposé de prendre

$$\forall h \in \{0,1,...,T\} \left\{ \begin{array}{ccc} k_h^{(out)} & = & |B - I_h| \\ k_h^{(in)} & = & \alpha \\ k_h^{(b)} & = & \lambda \end{array} \right.$$

L'énergie à minimiser est donc (Ω est indexé par h car la segmentation se fait sur chaque image I_h)

$$E_{ABJB}(\Omega_h) = \int_{\Omega_h} \alpha \, d\mathbf{x} + \int_{\Omega_h^c} |B - I_h|(\mathbf{x}) \, d\mathbf{x} + \lambda \int_{\Gamma_h} d\mathcal{H}^1$$

et l'EDP obtenue par descente de gradient de cette énergie est

$$\frac{\partial \Gamma_h}{\partial t} = -(\alpha - |B - I_h| + \lambda\kappa)\mathbf{n},$$

ce qui en termes d'ensembles de niveau s'écrit (u_h est initialisée comme la fonction distance à Γ, nous suivons la méthode employée par T. Chan et L. Vese dans [CV1])

$$\frac{\partial u_h}{\partial t} = \delta_\epsilon(u_h) \left(|B - I_h| - \alpha + \lambda \mathrm{div}\left(\frac{\nabla u_h}{|\nabla u_h|} \right) \right).$$

Il s'agit donc d'un *snake* image par image : sur chaque image I_h, on fait évoluer le contour initial afin de minimiser l'énergie correspondant aux descripteurs écrits ci-dessus.
A condition de bien ajuster le paramètre α, les zones telles que $\alpha^2 - (B - I_h)^2 < 0$ sont incluses dans les objets en mouvement. Il se trouve que si l'on dispose du résultat de l'algorithme d'Aubert, à savoir $(C_1, ..., C_h)$ et B, une valeur naturelle

pour α serait $\min_{(x,y)}[(B-I_h)^2(x,y)\cdot(1-C_h)(x,y)]$. Ou encore, sans effectuer ce calcul qui peut être plus ou moins coûteux, on peut considérer le cas $\alpha_c^r = 0$. L'énergie se simplifie alors en

$$E(B,C_1,...,,C_T) = \sum_{h=1}^{T}\int_\Omega C_h^2(B-I_h)^2\,dx + \alpha_c\sum_{h=1}^{T}\int_\Omega (C_h-1)^2\,dx + \alpha_b^r\int_\Omega \phi_1(|\nabla B|)\,dx$$

et l'équation d'Euler-Lagrange $(\nabla_{C_h}E, \delta C_h) = 0$ donne

$$\forall \delta C_h,\ \sum_{h=1}^{T}\int_\Omega C_h(B-I_h)^2\,\delta C_h\,dx + \alpha_c\sum_{h=1}^{T}\int_\Omega (C_h-1)\,\delta C_h\,dx = 0$$

on a alors nécessairement

$$C_h(B-I_h)^2 + \alpha_c(C_h-1) = 0$$

et finalement

$$C_h = \frac{\alpha_c}{\alpha_c + (B-I_h)^2}$$

En fonction de α_c et du seuil s que l'on choisit pour décider si C_h vaut 0 ou 1, on a ainsi une valeur explicite

$$\alpha_0^2 = \alpha_c\frac{1-s}{s}$$

Dans le cas où $\alpha_c^r \neq 0$, on peut chercher la bonne valeur de α autour de la valeur α_0. Ainsi on s'assure de bien avoir $\alpha - (B-I_h)^2 < 0$ dans les objets en mouvement.

On présente **en figure 3.8** des résultats obtenus en faisant varier le paramètre λ, on constate l'effet attendu, à savoir que plus λ est grand, plus les contours obtenus sont réguliers, et réciproquement plus λ est faible, plus les contours sont irréguliers et plus du bruit parasite est présent. Afin de montrer l'influence du paramètre λ, on présente le résultat du seuillage $|B-I_h| > \alpha$ qui correspond à la solution du problème de minimisation pour $\lambda = 0$. Finalement l'intérêt de cette segmentation par rapport à celle d'Aubert, Deriche et Kornprobst est de chercher à régulariser le résultat du seuillage $|B-I_h| > \alpha$ à l'aide du paramètre λ.

3.2.1.2 Formulation discrète, champs de Markov, ICM

On peut aussi implémenter un algorithme de minimisation de l'énergie

$$E(\Omega) = \alpha\int_\Omega dx + \int_{D\setminus\Omega}|B-I_h|\,dx + \lambda\int_{\partial\Omega} d\mathcal{H}^1$$

en réécrivant cette énergie sous forme discrète. Soit en effet l'énergie discrète

$$E(\phi) = \sum_{i,j}(\alpha - |B - I_h|(i,j))H(\phi(i,j)) + \lambda \sum_{i,j} \underbrace{\sqrt{(H(\phi(i+1,j)) - H(\phi(i,j)))^2 + (H(\phi(i,j+1)) - H(\phi(i,j)))^2}}_{\triangleq l(i,j)}.$$

Ici la variable ϕ vaut 1 si le pixel (i,j) appartient à la zone mobile et -1 sinon ; H désigne la fonction de Heaviside. Analysons le deuxième terme qui est l'analogue de la longueur en discret. Pour cela, il nous faut étudier la valeur de ϕ en (i,j) et en ses deux voisins $(i+1,j)$ et $(i,j+1)$.

- Si $\phi(i,j) = \phi(i+1,j) = \phi(i,j+1)$, alors $l(i,j) = 0$,
- si $\phi(i,j) = \phi(i+1,j) = -\phi(i,j+1)$ ou $\phi(i,j) = \phi(i,j+1) = -\phi(i+1,j)$, $l(i,j) = 1$,
- si $\phi(i,j) = -\phi(i,j+1) = -\phi(i+1,j)$, $l(i,j) = \sqrt{2}$.

L'algorithme pour minimiser l'énergie discrète est alors :

1. on balaye l'image (dans le sens naturel : horizontalement et verticalement),
2. on étudie la variation d'énergie

   $$\Delta E(i,j) = (\alpha - |B - I_h|(i,j)) * (1 - 2H(\phi(i,j))) + \lambda(\Delta l_{i,j} + \Delta l_{i-1,j} + \Delta l_{i,j+1})$$

   induite par $\phi(i,j) \longleftarrow -\phi(i,j)$. Si $\Delta E(i,j) < 0$, on change $\phi(i,j)$ en $-\phi(i,j)$, sinon on ne fait rien,
3. une fois le balayage effectué, on recommence les balayages jusqu'à convergence.

Les variations $\Delta l_{i,j}$, $\Delta l_{i-1,j}$, $\Delta l_{i,j-1}$ de la longueur induites par $\phi(i,j) \longleftarrow -\phi(i,j)$ sont définies par

$$\Delta l_{i,j} = \sqrt{(H(\phi(i+1,j)) - H(-\phi(i,j)))^2 + (H(\phi(i,j+1)) - H(-\phi(i,j)))^2} - \sqrt{(H(\phi(i+1,j)) - H(\phi(i,j)))^2 + (H(\phi(i,j+1)) - H(\phi(i,j)))^2}$$

$$\Delta l_{i-1,j} = \sqrt{(H(\phi(i-1,j+1)) - H(\phi(i-1,j)))^2 + (H(-\phi(i,j)) - H(\phi(i-1,j)))^2} - \sqrt{(H(\phi(i-1,j+1)) - H(\phi(i-1,j)))^2 + (H(\phi(i,j)) - H(\phi(i-1,j)))^2}$$

$$\Delta l_{i,j-1} = \sqrt{(H(-\phi(i,j)) - H(\phi(i,j-1)))^2 + (H(\phi(i+1,j-1)) - H(\phi(i,j-1)))^2} - \sqrt{(H(\phi(i,j)) - H(\phi(i,j-1)))^2 + (H(\phi(i+1,j-1)) - H(\phi(i,j-1)))^2}$$

En fait l'énergie dicrète peut s'interpréter comme l'énergie associée à un champ markovien Φ conditionnellement aux observations B, I_h et au paramètre α tel que (on note $d = (B, I_h, \alpha)$ et $\mathcal{D}$ la variable aléatoire dont d est la réalisation)

$$P(\Phi = \phi | \mathcal{D} = d) = \frac{\exp(-E(\phi | \mathcal{D} = d))}{Z}.$$

En effet, E peut s'exprimer sous la forme

$$E(\phi|d) = \sum_{c \in \mathcal{C}_1} E_{\mathcal{C}_1}(c) + \lambda \sum_{c \in \mathcal{C}_3} E_{\mathcal{C}_3}(c)$$

où

- $\mathcal{C}_1$ désigne l'ensemble des cliques (ensemble de points tous voisins) d'ordre 1 (l'ensemble est de cardinal 1) en (i, j), il n'y a donc que le pixel (i, j),
- $\mathcal{C}_3$ désigne l'ensemble des cliques d'ordre 3 en (i, j) qui sont au nombre de 4, seule la clique $\{(i, j), (i + 1, j), (i, j + 1)\}$ admet une énergie $E_{\mathcal{C}_3}$ non nulle en général (***cf.*** **figure 3.7**).

A partir de là, il est facile de montrer en utilisant le théorème de Hammersley-Clifford ([Ge]) que le champ aléatoire Φ est markovien pour le système de voisinage défini par les cliques d'ordre au plus égal à 3. Quand à l'algorithme discret que l'on vient de décrire, ce n'est rien d'autre que l'ICM (Iterated Conditional Modes) introduit par Besag dans [Be] et qui consiste à maximiser la probabilité conditionnelle $P(\Phi_{i,j} = \phi(i, j) |\, \Phi_{-i,-j}, \mathcal{D}_{i,j})$ ou plus simplement à minimiser l'énergie conditionnelle $E(\Phi_{i,j} = \phi(i, j) |\, \Phi_{-i,-j} = \phi_{-i,-j}, \mathcal{D}_{i,j})$, où

$$\phi_{-i,-j} = \{\phi_{k,l}, (k, l) \neq (i, j)\} \text{ et } \Phi_{-i,-j} = \{\Phi_{k,l}, (k, l) \neq (i, j)\}$$

Le système de voisinage employé ici est la 4-connexité. Les cliques de 3 pixels participent au terme de longueur même si le terme $l(i, j)$ est non nul pour un seul d'entre eux (***cf.*** **figure 3.7**).

L'algorithme des modes conditionnels itérés avait été proposé par Besag comme une alternative au recuit simulé proposé par Geman et Geman dans [Ge] qui converge certes vers un minimum global, mais en un temps très long. En revanche, l'ICM est connu pour converger seulement vers un minimum local de l'énergie, l'algorithme peut donc rester bloqué dans le bassin d'un minimum local mais il a l'avantage d'être particulièrement rapide (3 à 5 balayages suffisent pour atteindre la convergence). Cela a été vérifié par Chan et Song ([SC]) dans le cas de la fonctionnelle de Mumford et Shah à deux régions (intérieur et extérieur), bien que l'équivalence avec l'ICM n'ait pas été notée. Nous avons testé les algorithmes sur une séquence de voitures[2]. Une étude fine des **figures 3.8 et 3.9** montre toutefois que si le résultat de l'ICM est très comparable avec le résultat de la minimisation par EDP de la fonctionnelle continue, les deux résultats comportent tout de même des différences (certaines voitures sont "coupées" en deux,

[2]librement téléchargeable sur le site web de l'université de Karlsruhe `i21www.ira.uka.de/image_sequences`

certains coins sont moins naturels dans le résultat de l'ICM). Pour bénéficier des avantages des deux algorithmes (rapidité pour l'ICM, précision pour la minimisation par EDP), on peut alors — comme le proposaient Chan et Song — effectuer l'algorithme discret, puis quelques itérations de l'algorithme par EDP. Cette méthode se montre en effet très efficace, car elle permet de débruiter davantage qu'en appliquant directement l'ICM tout en bénéficiant de la rapidité de cet algorithme, un nombre d'itérations relativement petit étant nécessaire pour l'algorithme de snakes. **Sur les figures 3.9, 3.10 et 3.11**, on montre les résultats obtenus successivement pour l'algorithme de snakes à partir de l'initialisation donnée par le seuillage $|B - I_h| > \alpha$ après 1000 et 3000 itérations, puis pour l'algorithme de snakes précédé de l'ICM après seulement 1000 itérations. On voit **en figures 3.10 et 3.11** que le résultat est visuellement meilleur (il y a moins de bruit) lorsque l'ICM est appliqué avant pour un même nombre d'itérations (ici 1000 et 3000) de l'algorithme de snakes, ce qui est confirmé par la différence entre les masques montrée **en figure 3.12**.

3.2.1.3 ICM vs EDP : temps de calcul et résultats

Comme nous l'avons dit, l'ICM peut être appliqué avant l'algorithme de minimisation de la fonctionnelle continue par EDP pour réduire le temps de calcul de l'EDP. Afin de vérifier ceci, nous avons poussé les algorithmes jusqu'à convergence. C'est-à-dire que l'on stoppe l'ICM lorsque la différence entre les masques de chaque début de balayage est nulle. Le critère d'arrêt pour l'EDP est un peu plus compliqué ; on stoppe l'algorithme lorsque l'approximation de l'aire de la région Ω évolue de moins de 1 pixel :

$$A^\epsilon(u^n) = \sum_{s \in D} H_\epsilon(u^n),$$

le critère d'arrêt est donc $|A^\epsilon(u^{n+1}) - A^\epsilon(u^n)| < \eta$ où $\eta < 1$.
Le tableau suivant donne une idée des temps de calcul lorsque l'on choisit l'initialisation $\{|B - I| \leq \alpha\}$ ($\eta = 0,7$, pas de temps de l'algorithme $\delta t = 5$, $\alpha = 20$, dixième image de la séquence). On vérifie l'idée intuitive selon laquelle l'algorithme devrait être plus lent au fur et à mesure que l'on augmente λ. Cela est toutefois assez peu vérifié dans le cas de l'ICM, mais la sensibilité aux minima locaux joue ici un rôle non négligeable. Enfin, on vérifie que l'application de l'ICM avant d'appliquer l'EDP permet de réduire le temps de calcul de l'EDP pour les plus grandes valeurs de λ. Les différents résultats (ICM, EDP, ICM+EDP) pour $\lambda = 20$ sont montrés en **figure 3.13**. Les algorithmes ont été appliqués sur un processeur Pentium 4 M de PC portable à $1,8$ GHz et 1 Go de RAM.

	Temps de calcul en secondes pour différentes valeurs de λ		
	ICM	EDP	ICM+EDP
$\lambda = 5$	$1,39$	68	$72,39$
$\lambda = 10$	$1,41$	75	$84,41$
$\lambda = 15$	$1,36$	133	$80,36$
$\lambda = 20$	$1,38$	117	$58,38$
$\lambda = 25$	$1,36$	122	$49,36$

L'étude des différents temps de calcul ne doit pas nous faire oublier que l'implémentation par EDP de la minimisation de la fonctionnelle continue est fondamentalement différente de la minimisation de la fonctionnelle discrète par ICM sur un point. En effet, bien souvent, on se sert de la fonction

$$\delta_\epsilon(s) = \frac{2}{\pi}\frac{\epsilon}{\epsilon^2 + s^2}$$

pour approcher le dirac intervenant dans l'EDP

$$\frac{\partial u}{\partial t} = \delta(u)\left(|B - I| - \alpha + \lambda \text{div}\left(\frac{\nabla u}{|\nabla u|}\right)\right).$$

Or cette approximation n'est jamais nulle, elle a donc pour effet de faire intervenir les valeurs de u qui sont très éloignées de la ligne de niveau 0, ce qui est évidemment très différent du comportement de l'ICM où l'on ne fait intervenir que des pixels voisins de celui que l'on étudie. Toutefois nous avons choisi de garder cette approximation du dirac, car elle donne souvent de meilleurs résultats (*cf.* [CV1]).

3.2.1.4 Remarque sur le modèle d'Ising et la "longueur du chauffeur de taxi"

Le potentiel $E_{\mathcal{C}_3}$ se distingue du potentiel d'Ising qui ne fait intervenir que les cliques d'ordre 2. En effet, le potentiel d'Ising est donné (avec les notations que nous avons utilisé) par

$$|H(\phi(i+1,j)) - H(\phi(i,j))| + |H(\phi(i,j+1)) - H(\phi(i,j))|.$$

On voit que cette quantité est égale à

$$\| \left(H(\phi(i+1,j)) - H(\phi(i,j)), H(\phi(i,j+1)) - H(\phi(i,j))\right) \|_1$$

qui est la discrétisation de

$$\|(\partial_x(H(\phi)), \partial_y(H(\phi)))\|_1 = \|\nabla H(\phi)\|_1.$$

Le modèle d'Ising résulte donc lui aussi d'une discrétisation, mais d'une longueur "pondérée". En effet, en description continue, on aurait $\int_D \|\nabla H(\phi)\|_1$ et

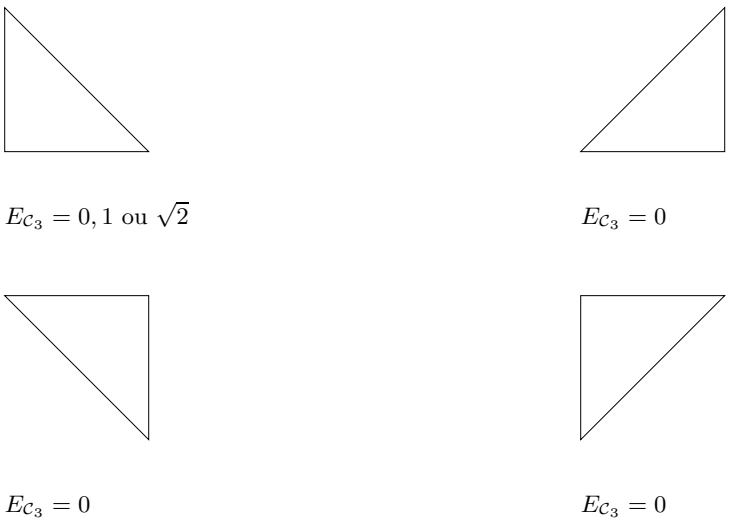


FIG. 3.7: Les différents systèmes de voisinage de trois pixels et leurs énergies associées

non $\int_D \|\nabla H(\phi)\|_2$ qui est la longueur au sens de Hausdorff. Toutefois, on peut prouver (*cf.* [Cham]) que les mesures $\|\nabla H(\phi)\|_1$ et $\|\nabla H(\phi)\|_2$ sont absolument continues l'une par rapport à l'autre, et donc on peut écrire

$$\Lambda(\Gamma) \triangleq \int_D \left\| \frac{\nabla \phi}{\|\nabla \phi\|_2} \right\|_1 \|\nabla H(\phi)\|_2 = \int_D \|\nu\|_1 \|\nabla H(\phi)\|_2 = \int_\Gamma \|\nu\|_1 \, d\mathcal{H}^1,$$

où ν désigne la normale (définie presque partout au sens de la mesure de Hausdorff) à la courbe Γ. On voit que Λ se distingue de la mesure de Hausdorff

$$\mathcal{H}^1(\Gamma) = \int_D \|\nabla H(\phi)\|_2 = \int_\Gamma d\mathcal{H}^1.$$

Elle est appelée "longueur du chauffeur de taxi" par Chambolle, en référence au trajet que doit parcourir un taxi dans une ville dont le réseau est formé d'une grille de rues perpendiculaires. Pour une étude théorique de la métrique du chauffeur de taxi appliquée à la fonctionnelle de Mumford et Shah, on se référera à [Cham].

3.2.2 Suivi d'objets en mouvement : Une méthode de tracking sans modèle de mouvement *a priori*

Une autre approche qui n'est pas de la segmentation d'objets en mouvement au sens premier du terme est de segmenter en premier lieu un objet dont on sait qu'il est mouvement et de chercher à le suivre tout au long de la séquence. Si on voulait faire cela de manière automatique, sans connaissance *a priori* sur le type de séquence d'images que l'on traite, ce serait fastidieux, car il nous faudrait

FIG. 3.8: En haut, l'image du fond et la segmentation obtenue par le seuillage $|B-I_h| > \alpha$, au milieu et à droite : segmentation par l'algorithme Aubert–Barlaud–Jehan-Besson pour $\lambda = 10$ et $\lambda = 20$ ($\delta t = 5$, $\epsilon = 1,5$, $\alpha = 20$, l'initialisation choisie est montrée sur l'image du bas). Le paramètre λ permet de régulariser le résultat du seuillage $|B-I_h| > \alpha$.

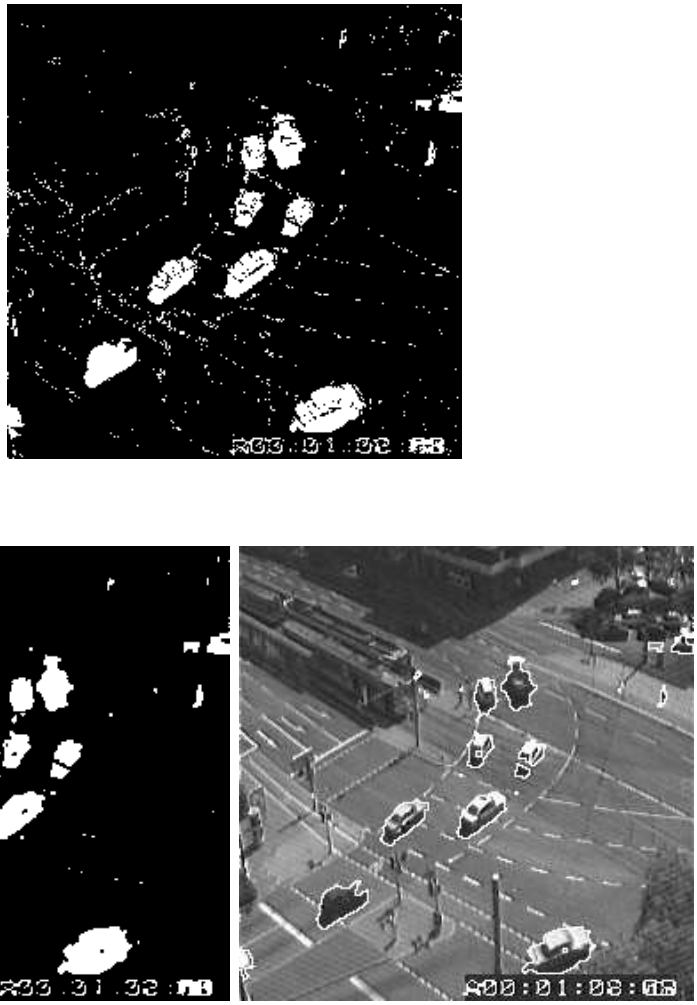


FIG. 3.9: En haut, le masque obtenu par seuillage de $B - I_h$ au niveau α qui sert d'initialisation aux algorithmes, puis le masque obtenu en lui appliquant l'ICM (4 balayages, $\alpha = 20$, $\lambda = 10$).

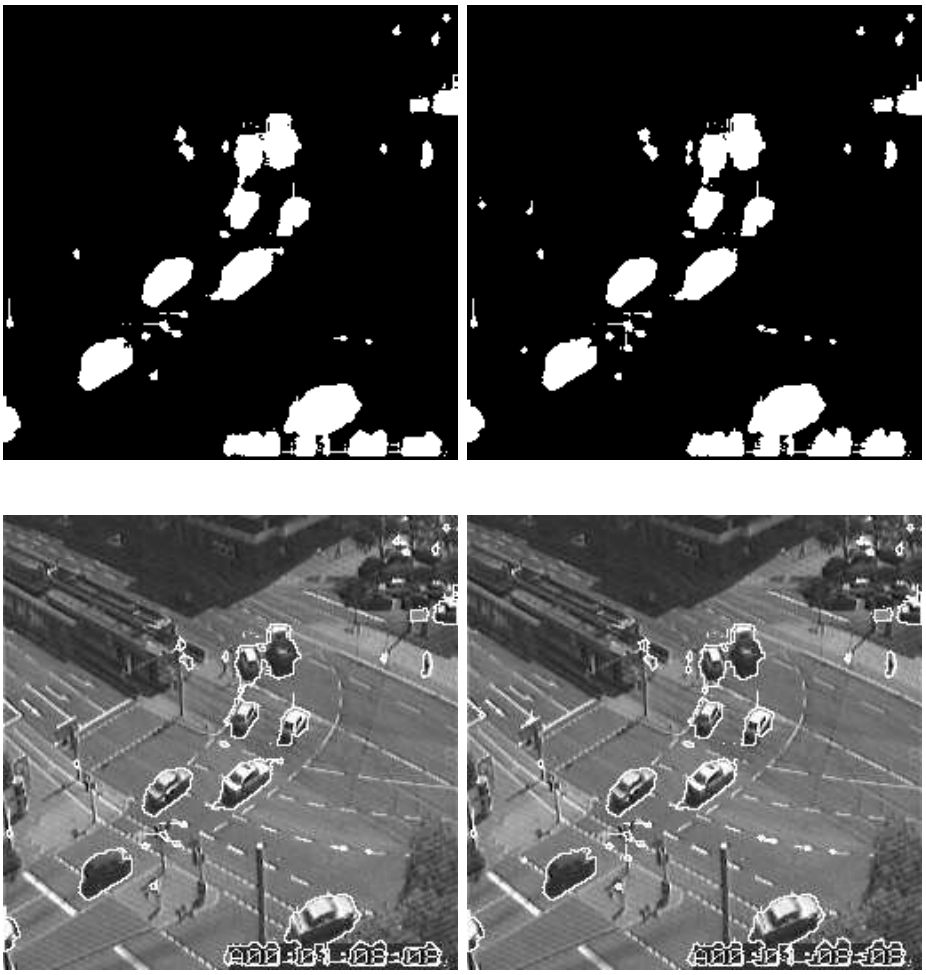


FIG. 3.10: Suite de la figure 3.9. On reprend comme initialisation le masque obtenu par seuillage de $B - I_h$ au niveau α, montré sur la figure 3.9. De gauche à droite et de haut en bas : les masques obtenus en combinant ICM (4 balayages) et EDP sur l'algorithme Aubert–Barlaud–Jehan-Besson ($\epsilon = 1.$, $\lambda = 10$, 1000 itérations), puis ceux en utilisant seulement l'EDP (mêmes paramètres). En bas les masques superposés aux images correspondantes. On obtient un résultat meilleur au prix du très léger temps de calcul de l'ICM en appliquant l'ICM avant l'algorithme par EDP.

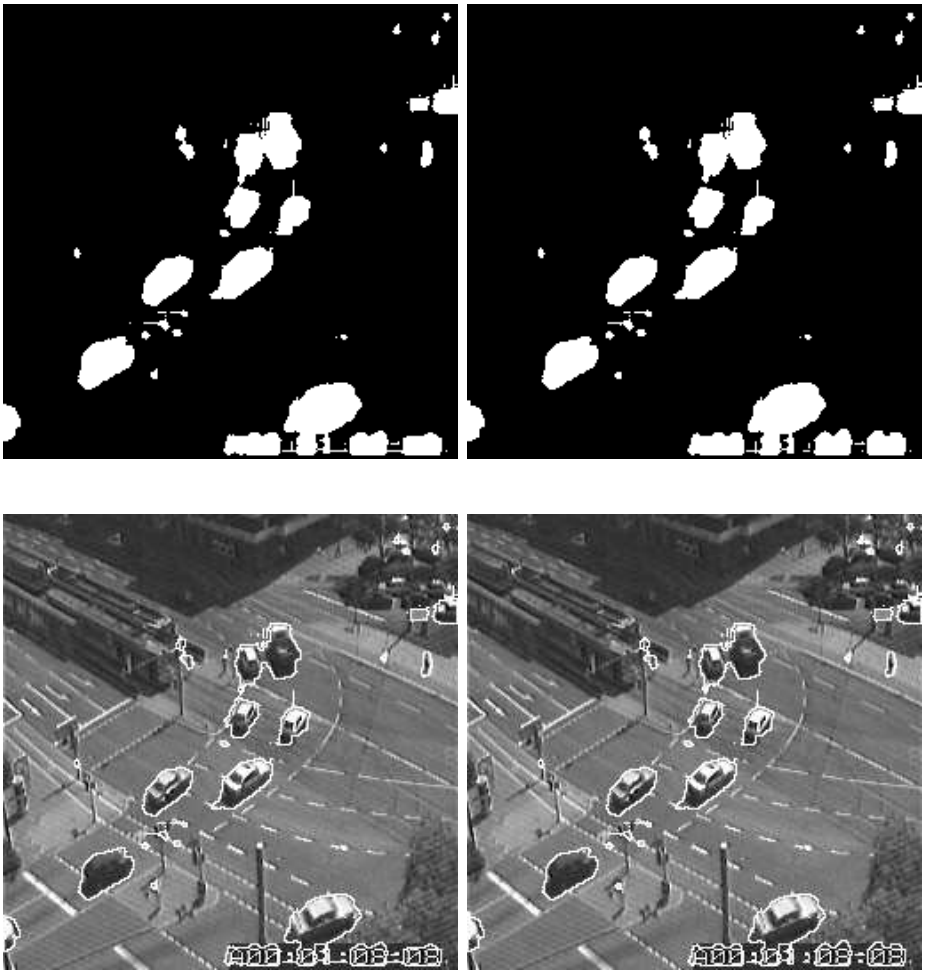


FIG. 3.11: Suite de la figure 3.10. En haut de gauche à droite : les masques obtenus en combinant ICM (4 balayages) et EDP sur l'algorithme Aubert–Barlaud–Jehan-Besson ($\epsilon = 1$, $\lambda = 10$, 3000 itérations), puis ceux en utilisant seulement l'EDP (mêmes paramètres). En bas les masques superposés aux images correspondantes. On obtient un résultat meilleur (comparer avec la figure précédente) au prix du très léger temps de calcul de l'ICM en appliquant l'ICM avant l'algorithme par EDP.

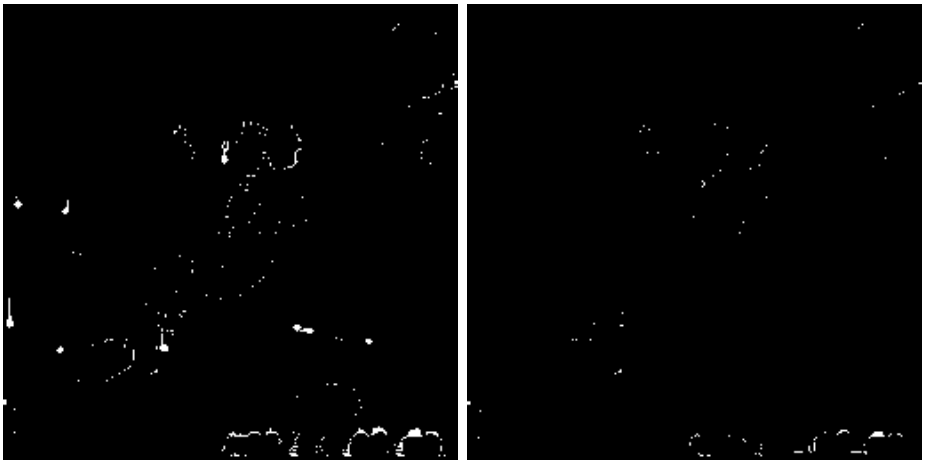


FIG. 3.12: Suite des figures 3.10 et 3.11. Différence entre les masques obtenus avec l'ICM+EDP et l'EDP (algorithme Aubert–Barlaud–Jehan-Besson) pour les masques des figures 3.10 et 3.11. Ces différences montrent le bruit entre les deux résultats.

détecter les objets en mouvement qui apparaissent ou qui disparaissent pendant la séquence, qui se mettent à bouger alors qu'ils étaient immobiles... Pour cette raison, le *suivi* d'objets en mouvement ou *tracking* et la segmentation au sens de la *détection* d'objets en mouvement sont deux approches qui répondent à deux buts différents. Comme dans la section précédente, nous allons présenter une méthode de tracking proposée par Mansouri, Mitiche et Aron dans [MMA], à la fois sous sa forme *snake*, puis sous la forme ICM.

3.2.2.1 Initialisation

Dans tout problème de tracking, se pose le problème de la segmentation dans l'image initiale de l'objet à suivre le long de la séquence. Cela peut se faire par une méthode de segmentation d'image quelconque ou bien à la main. Nous avons choisi d'utiliser une méthode dont l'implémentation était assez facile d'utilisation pour ne pas nous focaliser sur le problème de l'initialisation qui conditionne certes la qualité du tracking, mais qui doit se faire assez rapidement.
Pour cette raison, nous avons choisi d'appliquer une segmentation basée sur le contraste moyen dûe à Kimmel et Bruckstein ([KB]) et reprise dans [DMM] pour une comparaison avec un détecteur de contours statistique. L'énergie à maximiser est

$$J_{KB}(\Gamma) = \frac{\int_\Gamma g(\nabla I \cdot \mathbf{n}_\Gamma)\, d\mathcal{H}^1}{\int_\Gamma d\mathcal{H}^1}$$

où g est en général choisi de la forme $g(t) = |t|^\alpha$. Pour le calcul de sa dérivée, on se reportera à l'annexe A.

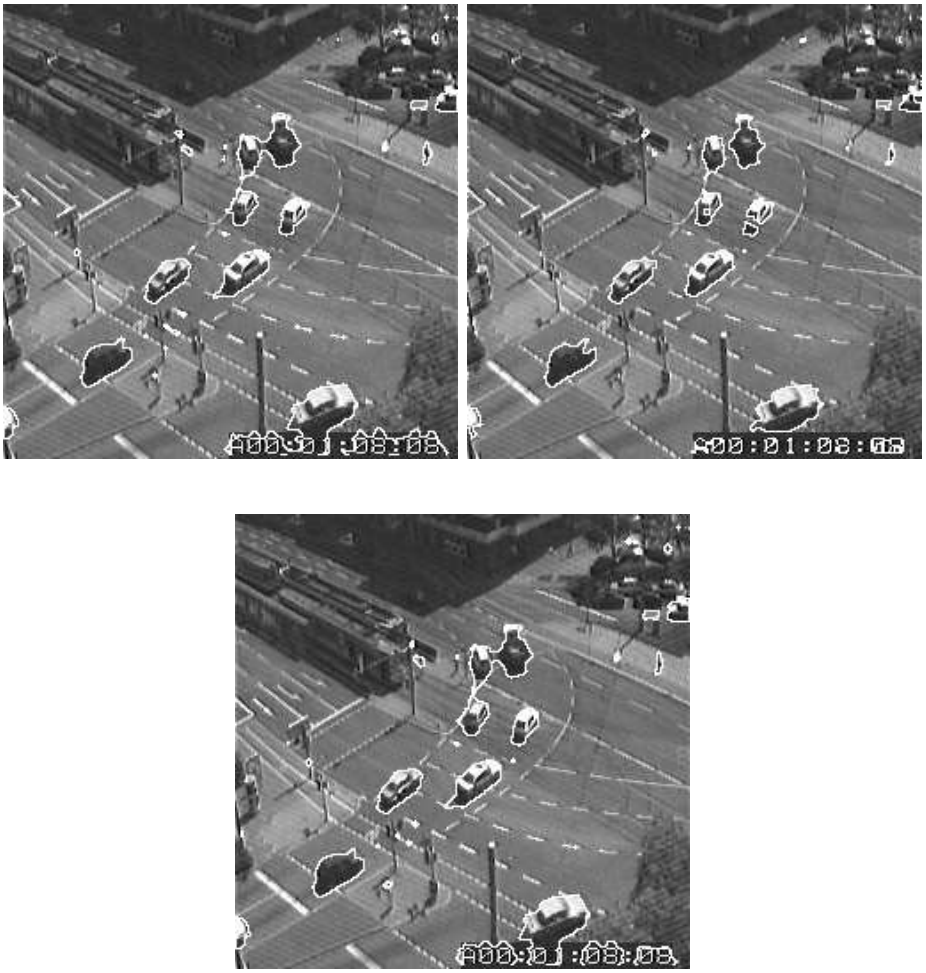


FIG. 3.13: De gauche à droite, résultats par application de l'EDP seule, de l'ICM seul, puis de l'ensemble des deux. Les paramètres sont $\epsilon = 1$, $\alpha = 20$, $\lambda = 20$, $\delta t = 5$ pour le pas de temps.

L'implémentation de L. Moisan dont nous disposons est une implémentation géométrique qui est assez sensible aux maxima locaux, il faut donc choisir le contour initial de façon à ce qu'il soit très proche de l'objet à segmenter. Sous ces conditions, le résultat de la segmentation sera correct.

3.2.2.2 Modèle continu

Dans [MMA], Mansouri, Mitiche et Aron proposèrent une méthode de tracking sans faire d'hypothèse sur le mouvement. Bien souvent, on fait l'hypothèse que le mouvement est paramétrique, ou bien on détermine un flot optique par morceaux par une méthode différentielle (Nagel-Enkelmann, Weickert-Schnörr,...). Ici, la seule hypothèse que les auteurs de [MMA] font est que le contour évolue dans un voisinage fixé autour du contour initial. En effet, une fois un paramètre α fixé, la région de référence Ω_{ref} et l'image de référence I_{ref} données, ils introduisent

$$\xi_1(\mathbf{x}) = \inf_{|v|\leq\alpha, \mathbf{x}+\mathbf{v}\in\Omega_{ref}} |I(\mathbf{x}) - I_{ref}(\mathbf{x}+\mathbf{v})|^2$$

et

$$\xi_2(\mathbf{x}) = \inf_{|v|\leq\alpha, \mathbf{x}+\mathbf{v}\in\Omega^c_{ref}} |I(\mathbf{x}) - I_{ref}(\mathbf{x}+\mathbf{v})|^2.$$

Lorsque l'on travaille sur une séquence d'images, I_{ref} est l'image précédente et Ω_{ref} le contour final trouvé dans l'image précédente. L'algorithme est ainsi appliqué de proche en proche sur les images de la séquence. Mais on remarque que ξ_1 n'est défini que sur le dilaté de Ω_{ref} (il faut qu'il existe $\mathbf{v}$ de norme inférieure à α, tel que $\mathbf{x}+\mathbf{v} \in \Omega_{ref}$) et que ξ_2 n'est défini que sur le dilaté du complémentaire de Ω_{ref}. L'intersection de ces deux ensembles est donc une bande autour de la frontière de Ω_{ref}, ainsi que le montre la figure 3.14.
Pour reprendre la terminologie des descripteurs de Barlaud *et al.*, les descripteurs choisis par Mansouri *et al.* sont

$$\begin{aligned} k_h^{(out)} &= \xi_2 \\ k_h^{(in)} &= \xi_1 \\ k_h^{(b)} &= \lambda. \end{aligned}$$

L'énergie à minimiser est donc

$$E_{MMA}(\Omega) = \int_\Omega \xi_1(\mathbf{x})\, d\mathbf{x} + \int_{\Omega^c} \xi_2(\mathbf{x})\, d\mathbf{x} + \lambda \int_{\partial\Omega} d\mathcal{H}^1$$

et l'EDP obtenue par descente de gradient de cette énergie est

$$\frac{\partial\Gamma}{\partial t} = -(\xi_1 - \xi_2 + \lambda\kappa)\mathbf{n},$$

ce qui en termes d'ensembles de niveau s'écrit (u est initialisée comme la fonction distance à Γ)

$$\frac{\partial u}{\partial t} = \delta_\epsilon(u)(\xi_2 - \xi_1 + \lambda \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)).$$

Implicitement, le calcul de ξ_1 et ξ_2 est un calcul de flot optique : on cherche le point $\mathbf{y}$ appartenant à la région de référence et à la boule de centre $\mathbf{x}$ et de rayon α tel que l'écart entre $I(\mathbf{x})$ et $I_{ref}(\mathbf{y})$ soit minimal. Cet écart est une mesure de l'erreur commise. Le choix des descripteurs impose que l'erreur relative à la région de référence soit minimale sur la région cherchée, de même pour son complémentaire.
Du fait que ξ_1 et ξ_2 ne sont pas définis sur tout le domaine mais sur la bande centrée sur le contour de référence et de rayon α (au sens de la dilatation et de l'erosion), le problème de minimisation est contraint à cette bande, et le contour solution appartient nécessairement à la bande. On peut aussi définir ξ_1 et ξ_2 sur tout le domaine en leur mettant la valeur $+\infty$ lorsqu'ils sont indéfinis.

3.2.2.3 Modèle discret

Comme le modèle présenté par Mansouri, Mitiche et Aron était présenté sous forme continu et dérivé pour obtenir une EDP, il nous a paru intéressant de l'écrire sous forme discrète et d'appliquer l'ICM, comme nous l'avons fait précédemment pour le modèle de segmentation d'Aubert *et al.*.
L'énergie discrète est donc

$$E_{MMA}(\phi) = \sum_{(i,j)\in B} (\xi_1(i,j) - \xi_2(i,j))H(\phi(i,j))$$
$$+ \lambda \sum_{(i,j)\in B} \underbrace{\sqrt{(H(\phi(i+1,j)) - H(\phi(i,j)))^2 + (H(\phi(i,j+1)) - H(\phi(i,j)))^2}}_{\triangleq l(i,j)}.$$

où B est la zone délimitée par la bande.
Le calcul à effectuer pour l'ICM est exactement le même que pour le modèle d'Aubert *et al.*, à ceci près que $\alpha - |B - I_h|$ est remplacé par $\xi_1 - \xi_2$.

Nous présentons des résultats **en figures 3.15, 3.16, 3.17 et 3.18**. Ces résultats illustrent bien les avantages et les inconvénients de cette méthode : si l'objet à suivre est assez grand, assez bien contrasté, plutôt rigide, le tracking est effectué correctement, en revanche, on voit que le tracking n'est pas robuste aux occultations, puisque lorsque la voiture passe derrière le poteau, celui-ci agit comme une barrière sur le snake. La non-robustesse aux occultations est une conséquence du phénomène d'accumulation des erreurs inhérent à cette méthode. En effet, le seul contour de référence fiable que nous imposons se trouve sur la première image ; ensuite, pour une image donnée, le contour de référence est celui que nous avons calculé sur l'image précédente. De fait, si la donnée du contour est de moins en moins fiable, les erreurs engendrées sont de plus en plus grandes car l'algorithme n'a plus comme référence le contour initialement choisi sur la première image. Nous illustrons ce phénomène sur les **figures 3.19, 3.20, 3.21, 3.22** : la qualité du tracking est correcte sur les premières images, mais les erreurs accumulées finissent par l'emporter.
Afin de corriger ces erreurs, nous avons essayé d'appliquer un algorithme de

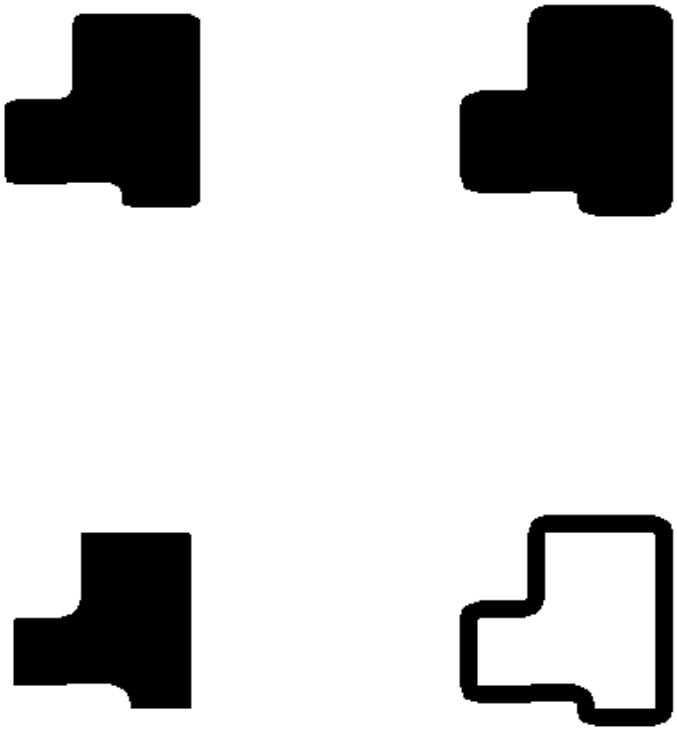


FIG. 3.14: Successivement, de haut en bas et de gauche à droite : la région de référence (en noir), sa dilatée (ensemble des points pour lesquels ξ_1 est définie), son érodée (ensemble des points pour lesquels ξ_2 est définie), et la bande (ensemble des points pour lesquels ξ_1 et ξ_2 sont tous deux définis).

contours actifs basé sur le niveau de gris. Sur chaque image, on applique l'algorithme de Mansouri, Mitiche et Aron pour obtenir un contour, puis on fait évoluer ce contour avec l'algorithme de Kimmel et Bruckstein. On rappelle (*cf.* section 3.2.2.1) que l'algorithme de Kimmel et Bruckstein est basé sur la maximisation de

$$J_{KB}(\Gamma) = \frac{\int_\Gamma g(\nabla I \cdot \mathbf{n}_\Gamma)\, d\mathcal{H}^1}{\int_\Gamma d\mathcal{H}^1}.$$

Cette idée assez simple permet d'éviter d'utiliser le modèle 3D exposé dans la section suivante. Les images de droite des **figures 3.19, 3.20, 3.21, 3.22** montrent les résultats obtenus en appliquant cette méthode.

3.2.2.4 Avantages et inconvénients de la méthode de tracking 2D

Enfin, notons qu'il n'est pas aisé de trouver un algorithme de tracking qui soit robuste aux occultations. Pour ce faire, on pourrait introduire une contrainte de préservation de forme ou de rigidité dans le modèle énergétique (voir par exemple les travaux de Chen et Vemuri [CheV]), et ce à condition que l'objet occultant soit de faible taille par rapport à l'objet occulté : si une partie significative de l'objet est occulté, il sera difficile de reconstituer l'objet, l'algorithme essayant de compenser les différents termes de l'énergie. On peut aussi utiliser des modèles 3D. Par exemple, Mansouri, Mitiche et Aron avaient généralisé leur modèle 2D en un modèle de tracking spatio-temporel où l'inconnue est cette fois un volume V qui minimise

$$E_{MMA^2}(V) = \int_V \xi_1(\mathbf{x},t)\, d\mathbf{x}dt + \int_{V^c} \xi_2(\mathbf{x},t)\, d\mathbf{x}dt + \lambda \int_{\partial V} d\mathcal{H}^2$$

où $d\mathcal{H}^2$ représente la mesure de Hausdorff de dimension 2 dans l'espace de dimension 3 $D \times [0,T]$ et

$$\xi_1(\mathbf{x},t) = \inf_{|v|\leq t\alpha, \mathbf{x}+\mathbf{v}\in\Omega_{ref}} |I(\mathbf{x}) - I_{ref}(\mathbf{x}+\mathbf{v})|^2,$$

$$\xi_2(\mathbf{x},t) = \inf_{|v|\leq t\alpha, \mathbf{x}+\mathbf{v}\in\Omega^c_{ref}} |I(\mathbf{x}) - I_{ref}(\mathbf{x}+\mathbf{v})|^2,$$

avec cette fois-ci I_{ref} qui est la première image et Ω_{ref} une région donnée sur la première image. Par rapport au modèle 2D, Mansouri *et al.* avaient noté que le modèle 3D exposé ci-dessus évitait davantage le problème d'accumulation d'erreurs observé dans le cas 2D, toutefois cela nécessite des calculs plus longs et plus lourds. En effet, la taille du voisinage qui sert à définir les ξ_i devient de plus en plus grande au cours du temps, ce qui rend le calcul des ξ_i de plus en plus lourd. D'après A. Mitiche (communication personnelle), on peut alors faire du tracking par "paquets" : on initialise le contour sur la première image, on applique l'algorithme 3D sur n images, puis sur les suivantes en se servant du contour sur l'image n (on réinitialise le temps t), et on itère le procédé.

FIG. 3.15: Application de la méthode de tracking de Mansouri *et al.* par l'ICM. Images paires de la séquence de la 2^e à la 12^e : 5 balayages de l'ICM sont effectués, $\lambda = 10$ et $\alpha = 3$.

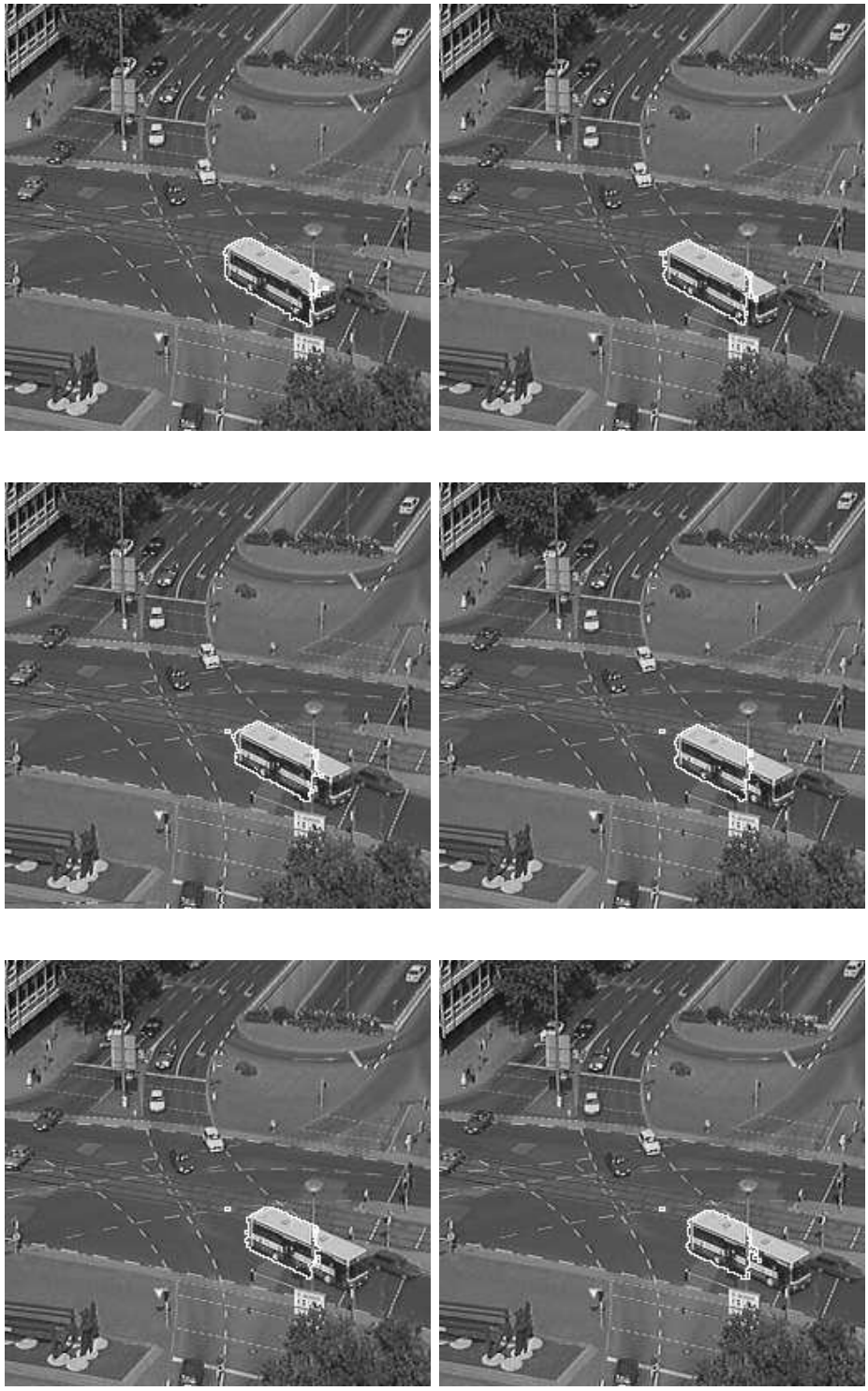


FIG. 3.16: Suite de la figure 3.15. Application de la méthode de tracking de Mansouri *et al.* par l'ICM. Images paires de la séquence de la 14^{e} à la 24^{e} : 5 balayages de l'ICM sont effectués, $\lambda = 10$ et $\alpha = 3$. L'algorithme n'est pas robuste aux occultations : le poteau occultant le bus empêche l'algorithme d'effectuer correctement le suivi.

FIG. 3.17: Suite de la figure 3.16. Application de la méthode de tracking de Mansouri *et al.* par l'ICM. Images paires de la séquence de la 26^{e} à la 36^{e} : 5 balayages de l'ICM sont effectués, $\lambda = 10$ et $\alpha = 3$. L'algorithme n'est pas robuste aux occultations : le poteau occultant le bus empêche l'algorithme d'effectuer correctement le suivi.

FIG. 3.18: Suite de la figure 3.17. Application de la méthode de tracking de Mansouri *et al.* par l'ICM. Images paires de la séquence de la 38^{e} à la 48^{e} : 5 balayages de l'ICM sont effectués, $\lambda = 10$ et $\alpha = 3$. Le suivi échoue complètement à partir du moment où le bus s'éloigne du poteau.

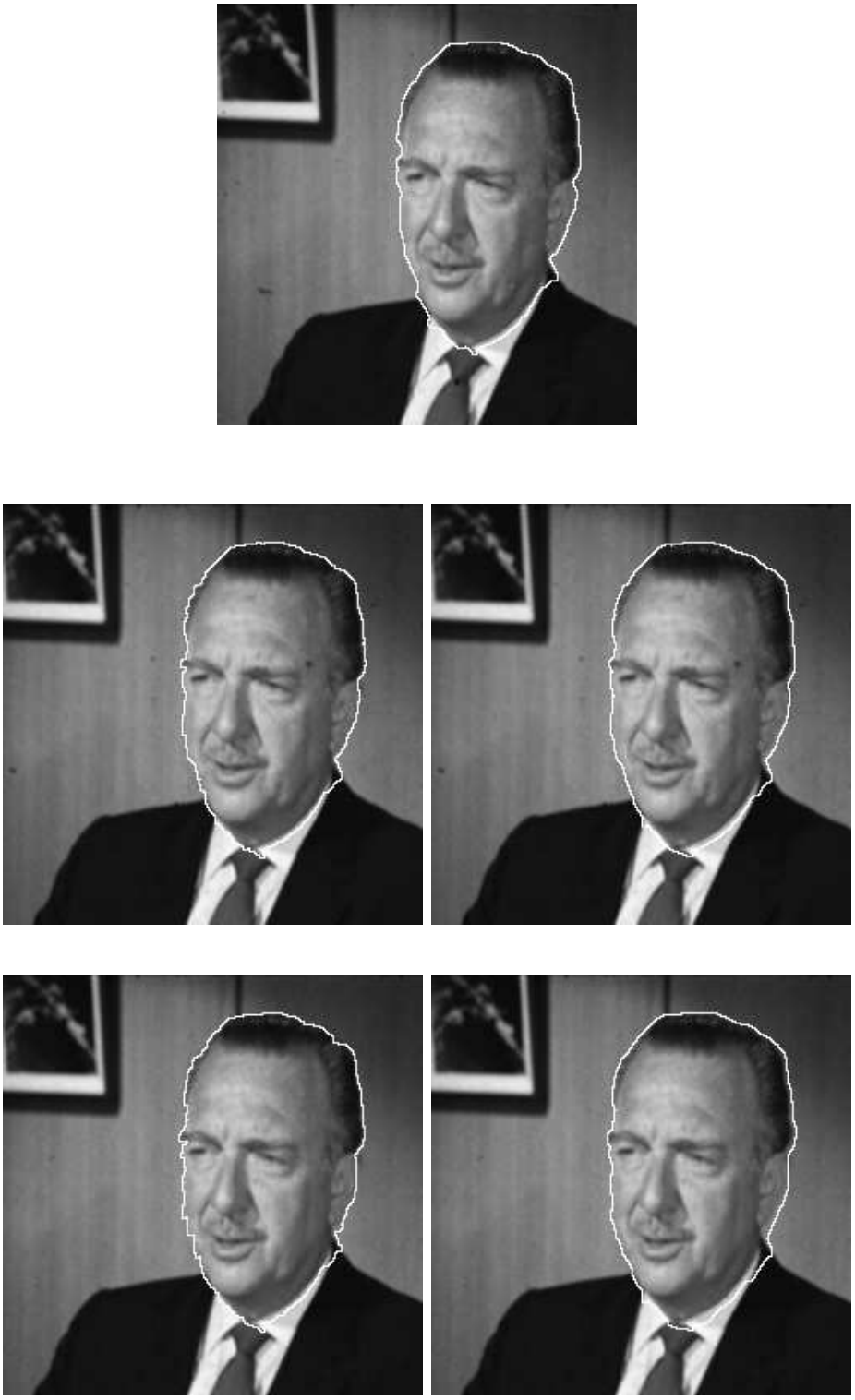


FIG. 3.19: En haut : contour de référence trouvé par la méthode de Kimmel-Bruckstein. Sur les images de gauche, résultat de la méthode de tracking de Mansouri *et al.* par l'ICM, puis le résultat obtenu par le même algorithme auquel on rajoute un contour actif basé niveau de gris de Kimmel-Bruckstein. Les résultats sont comparables car les erreurs ne sont que peu à peu amplifiées.

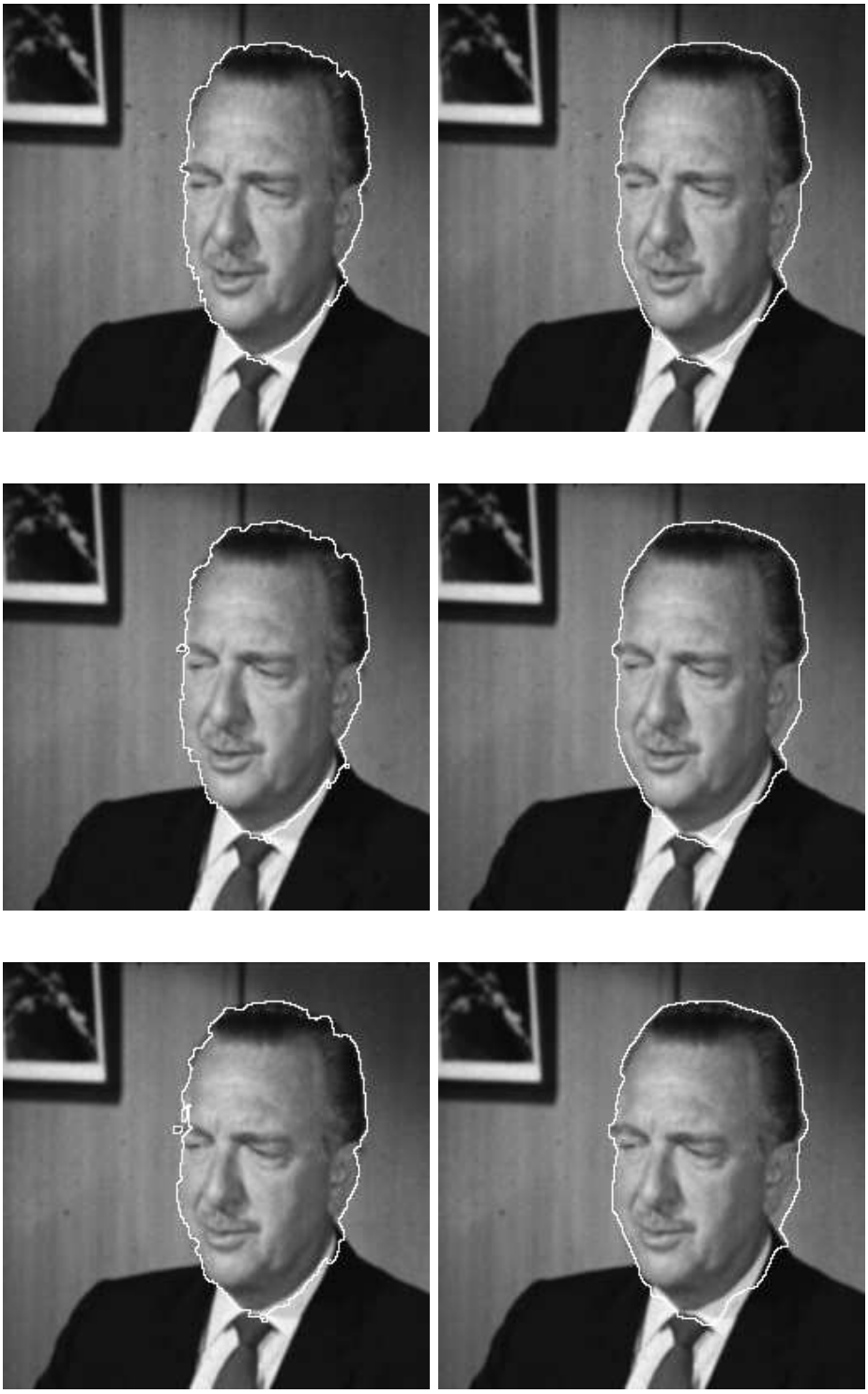


FIG. 3.20: Suite de la figure 3.19. Sur les images de gauche, résultat de la méthode de tracking de Mansouri *et al.* par l'ICM, puis le résultat obtenu par le même algorithme auquel on rajoute un contour actif basé niveau de gris de Kimmel-Bruckstein. Les résultats sont comparables car les erreurs ne sont que peu à peu amplifiées.

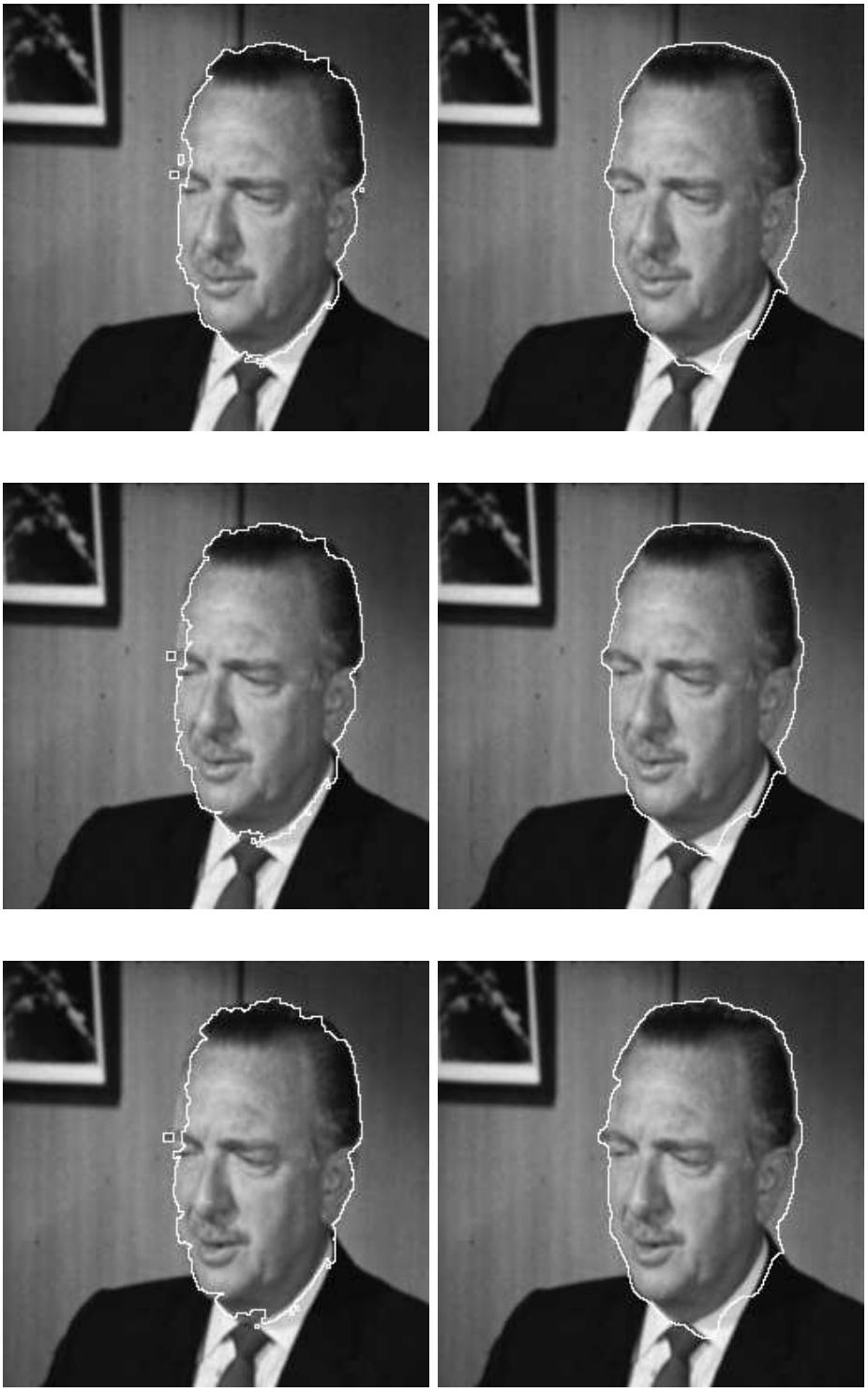


FIG. 3.21: Suite de la figure 3.20. Sur les images de gauche, résultat de la méthode de tracking de Mansouri *et al.* par l'ICM, puis le résultat obtenu par le même algorithme auquel on rajoute un contour actif basé niveau de gris de Kimmel-Bruckstein. Les différences commencent à apparaître. Le tracking par la méthode de Mansouri utilise le contour calculé sur l'image précédente. Comme celui-ci tend à s'éloigner des bords réels, les erreurs s'accumulent de plus en plus. La combinaison de l'algorithme avec un algorithme de snake en niveaux de gris permet d'éviter cela.

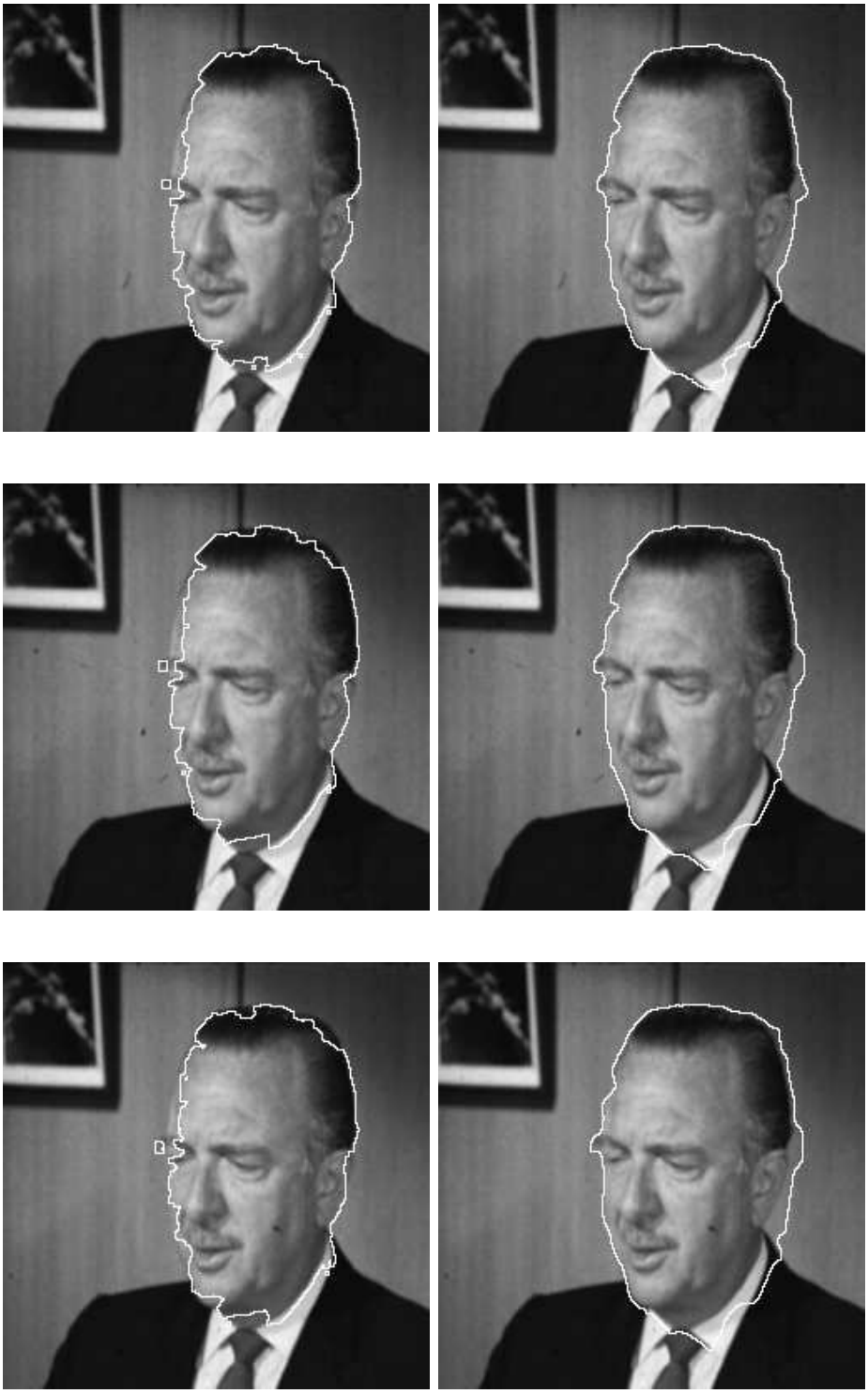


FIG. 3.22: Suite de la figure 3.21. Sur les images de gauche, résultat de la méthode de tracking de Mansouri *et al.* par l'ICM, puis le résultat obtenu par le même algorithme auquel on rajoute un contour actif basé niveau de gris de Kimmel-Bruckstein. Les différences sont de plus en plus manifestes.

3.3 Estimation du flot optique et segmentation d'objets en mouvement par utilisation du flot optique

Dans cette section, nous allons nous intéresser aux problèmes liés à l'estimation et à la segmentation du flot optique. Dans un premier temps, nous présenterons donc le problème de l'estimation et les améliorations que l'on peut y apporter, notamment la multirésolution. Ensuite, nous ferons un rapide état de l'art sur la segmentation du flot optique. Enfin, en utilisant les techniques de contours actifs que nous avons déjà présenté, nous présenterons un nouveau modèle.

3.3.1 Le problème du flot optique

En traitement d'images et en vision par ordinateur, l'outil couramment utilisé pour calculer le mouvement dans une séquence d'images est le flot optique. Le flot optique désigne la distribution des vitesses apparentes des motifs lumineux dans une prise de vue d'une caméra ou d'un capteur vidéo.
Il existe différentes modélisations mathématiques pour le calcul de flot optique, toutes celles que l'on considérera reposent sur le principe de conservation de l'intensité lumineuse I. Ainsi, si le point (x_1, x_2) suit une trajectoire $(t, x_1(t), x_2(t))$, on a

$$I(x_1(t), x_2(t), t) = cte.$$

On peut dériver cette relation par rapport au temps, cela nous donne

$$< \nabla I, \left(\frac{\mathrm{d}x_1}{\mathrm{d}t} \ \frac{\mathrm{d}x_2}{\mathrm{d}t}\right)^t > + \frac{\partial I}{\partial t} = 0. \tag{3.1}$$

Mais $\left(\frac{\mathrm{d}x_1}{\mathrm{d}t} \ \frac{\mathrm{d}x_2}{\mathrm{d}t}\right)^T$ représente le vecteur tangent à la trajectoire $(x_1(t), x_2(t))$, c'est à dire le vecteur vitesse $\mathbf{v}$. C'est cette quantité que nous définissons comme le flot optique.
Le flot optique est donc solution de (3.1), mais ce problème est mal posé, car le flot optique a deux composantes et (3.1) est la seule équation dont nous disposons. C'est le *problème d'ouverture*. Pour pouvoir le surmonter, il nous faut donc rajouter une contrainte sur le flot optique.
Dans le travail pionnier de Horn et Schunck ([HS]), le flot optique est supposé de norme H^1 finie, aussi proposent-ils de rechercher le flot optique comme la solution du problème de minimisation de la fonctionnelle

$$E_{HS}(v_1, v_2) = \int_D \alpha^2(|\nabla v_1|^2 + |\nabla v_2|^2) + (< \nabla I, \mathbf{v} > + \partial_t I)^2 \, d\mathbf{x}. \tag{3.2}$$

Si (v_1, v_2) minimise E, il vérifie les équations d'Euler-Lagrange

$$\begin{aligned} \alpha^2 \Delta v_1 &= I_x(< \nabla I, \mathbf{v} > + \partial_t I) \\ \alpha^2 \Delta v_2 &= I_y(< \nabla I, \mathbf{v} > + \partial_t I) \end{aligned}$$

Ces EDP s'apparentent donc à l'équation de la chaleur qui a l'inconvénient de lisser les discontinuités. De plus, l'hypothèse que le flot soit régulier au sens de la norme H^1 correspond mal au problème physique du flot optique que l'on voudrait lisser mais pas au voisinage des discontinuités de l'image. Afin de pouvoir effectivement réaliser ceci, plusieurs auteurs ont cherché d'autres fonctionnelles. On peut citer Nagel et Enkelmann ([NE]), Aubert, Deriche et Kornprobst ([ADK1]) et enfin Weickert et Schnörr ([WS]).
Nagel et Enkelmann ne linéarisent pas le terme de conservation de l'intensité : au lieu de $(< \nabla I_1, \mathbf{v} > + I_2 - I_1)^2$, ils écrivent $(I_2(\mathbf{x}+\mathbf{v}(\mathbf{x})) - I_1(\mathbf{x}))^2$. Quand au terme de régularisation, il est choisi de la forme

$$\int_D \frac{1}{|\nabla I_1| + 2\lambda^2} \mathrm{trace}((\nabla \mathbf{v})^T D(\nabla I_1)(\nabla \mathbf{v}))\, d\mathbf{x}$$

avec

$$D(\nabla I_1) = (\nabla^\perp I_1)(\nabla^\perp I_1)^T + \lambda^2 Id.$$

Leur idée est donc d'atténuer le lissage à travers les bords d'intensité $|\nabla I_1| \gg \lambda$.

Aubert, Deriche et Kornprobst choisissent de minimiser

$$E_{ADK}(\mathbf{v}) = \int_\Omega |\nabla I \cdot \mathbf{v} + \partial_t I| + \alpha \left[\int_\Omega \phi(\nabla v_1) + \int_\Omega \phi(\nabla v_2) \right] + \alpha_H \int_\Omega \frac{1}{1+|\nabla I|^2} |\mathbf{v}|_2^2 \, d\mathbf{x} \quad (3.3)$$

où ϕ est l'une des fonctions classiquement utilisée en traitement d'images pour préserver les discontinuités. Le dernier terme vise à atténuer le flot sur les zones de l'image à faible gradient, ce terme est essentiel pour les résultats d'existence et d'unicité, mais discutable sur le plan de la modélisation, selon P. Kornprobst ([Ko]).

Les approches présentées jusqu'ici reposent sur l'hypothèse que le flot admet une certaine régularité *spatiale*, c'est-à-dire sur un certain voisinage de chaque point de chaque image. Weickert et Schnörr font le même type d'hypothèse où le voisinage est cette fois étendu temporellement. La fonctionnelle à minimiser est alors

$$E(\mathbf{v}) = \int_{\Omega\times[0,T]} |\nabla I \cdot \mathbf{v} + \partial_t I|^2 \, d\mathbf{x}\, dt + \alpha \int_{\Omega\times[0,T]} \Psi(|\nabla^\theta v_1|^2 + |\nabla^\theta v_2|^2)\, d\mathbf{x}\, dt \quad (3.4)$$

où $\nabla^\theta = \left(\frac{\partial}{\partial x} \; \frac{\partial}{\partial y} \; \frac{\partial}{\partial t}\right)^T$, $\lambda > 0$ et $\Psi(s^2) = \epsilon s^2 + (1-\epsilon)\lambda^2\sqrt{1+\frac{s^2}{\lambda^2}}$.
Expérimentalement, l'hypothèse de régularité spatio-temporelle du flot optique permet de réduire considérablement la présence de bruit dans le flot optique tout en conservant bien les discontinuités du flot près des bords de l'image.

D'autres méthodes supposent le flot optique constant sur un voisinage $\mathcal{V}(\mathbf{x})$ de chaque point $\mathbf{x}$. Dans ce cas, si on cherche à minimiser $\int_{\mathcal{V}(\mathbf{x})}(< \nabla I, \mathbf{v} > +\partial_t I)^2 \, d\mathbf{x}$, le flot optique est solution du système linéaire

$$\left(\int_{\mathcal{V}(\mathbf{x})} (\nabla I)(\nabla I)^T(\mathbf{x}) \, d\mathbf{x} \right) \mathbf{v} = - \left(\int_{\mathcal{V}(\mathbf{x})} (\partial_t I \, \nabla I)(\mathbf{x}) \, d\mathbf{x} \right)$$

Cette approche est dûe à Lucas et Kanade ([LK]), et peut être formellement étendue à un voisinage spatio-temporel, mais en général, on choisit plutôt de chercher le vecteur propre pour la valeur propre minimale du tenseur

$$S(x) = \int_{\mathcal{V}(\mathbf{x})} (\nabla I)(\nabla I)^T(\mathbf{x}) \, d\mathbf{x}.$$

On renormalise quand c'est possible ce vecteur propre $\mathbf{e}_{min}$ afin que sa troisième composante soit 1. Le flot optique est alors

$$\mathbf{v} = (\mathbf{e}^1_{min} \; \mathbf{e}^2_{min})^T.$$

Cette méthode a été introduite par Bigün, Granlund et Wiklund dans [BGW]. La dernière remarque que nous pouvons faire est qu'il ne change rien dans tout ce qui précède de remplacer les intégrales sur $\mathcal{V}(\mathbf{x})$ par des convolutions par un noyau gaussien centré en $\mathbf{x}$ et d'écart-type fixé.

3.3.2 Remarque sur le flot optique avec multirésolution

La procédure de *multirésolution* a été introduite par Mémin et Pérez dans [MP] pour s'affranchir du problème de localité du flot optique. En effet, lorsque le déplacement est beaucoup plus important que le pixel, il sera peu vraisemblable que le flot optique corresponde à ce déplacement, puisque le calcul utilise le gradient de l'image qui est calculé en pratique sur un petit voisinage (3×3 ou 2×2) du pixel.
La procédure de multirésolution est la suivante :

1. On utilise une version I^J sous-échantillonnée de l'image I à la résolution J (la résolution la plus fine est 0). On minimise alors E pour trouver $\mathbf{v}^J$.
2. A la résolution $j < J$, on utilise le flot optique calculé à la résolution $j+1$ en cherchant le résidu $\mathbf{w}^j = \mathbf{v}^j - T\mathbf{v}^{j+1}$ (T est l'opérateur de sur-échantillonnage de la résolution $j+1$ vers la résolution j). On compense l'image par le mouvement $T\mathbf{v}^{j+1}$ dans la dérivée temporelle I^j_t et on remplace celle-ci par

$$\tilde{I^j_t}(\mathbf{x}, t) = I^j(\mathbf{x}, t+1) - I^j(\mathbf{x} - T\mathbf{v}^{j+1}, t).$$

3. La procédure est itérée de $j = J$ à $j = 0$.

En plus du problème de localité, cette technique permet aussi de mieux remplir les zones homogènes. En effet, dans une zone homogène, on a $|\nabla I| \approx 0$ et si cette zone homogène est très grande, il n'est pas évident que le flot soit important dans toute la zone, même s'il l'est sur sa frontière. Grâce à la technique de multirésolution, on étend plus rapidement le mouvement de la frontière à toute la zone.

3.3.3 Quelques modèles de segmentation par le flot optique

Une des premières segmentation du flot optique est dûe à Caselles et Coll ([CC]). Le flot optique est d'abord calculé par la méthode de Guichard (décomposition du flot optique suivant les deux directions ∇I et $(\nabla I)^{\perp}$) et ensuite les auteurs utilisent l'EDP des contours actifs géométriques à laquelle ils adjoignent le terme d'*advection* $< \mathbf{v}, \nabla u >$.

$$\frac{\partial u}{\partial t} = g\left(\operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) + \nu\right)|\nabla u| + (1-g) < \mathbf{v}, \nabla u > \tag{3.5}$$

En pratique, les auteurs choisissent plutôt de ne garder que le premier terme du membre de droite de (3.5) et transportent le contour final en connaissant la vitesse sur ce contour. Avec ces deux opérations, ils réussissent à faire du suivi d'objets éventuellement soumis à des occultations (*cf.* [CC]).

Une approche récente est dûe à Weickert *et al.* dans [KWSR]. Ils reprennent l'idée de Bigün *et al.* sur l'étude spectrale du tenseur de structure. Le voisinage est choisi assez grand (7^3 ou 11^3) pour réduire le bruit dans le flot optique. Mais l'idée des auteurs de [KWSR] est d'étudier si le flot optique peut être fiablement calculé, pour cela il introduit des mesures de cohérence

$$c_t = \begin{cases} 0 & \text{si } \lambda_1 = \lambda_3 \\ \exp(-\frac{\mu}{|\lambda_1 - \lambda_3|}) & \text{sinon} \end{cases}$$

et

$$c_s = \begin{cases} 0 & \text{si } \lambda_2 = \lambda_3 \\ \exp(-\frac{\mu}{|\lambda_2 - \lambda_3|}) & \text{sinon} \end{cases} .$$

Le flot optique est fiable si $c_t(x,y)$ est suffisamment grand et si le flot optique lui même est suffisamment grand en norme, ainsi, Weickert *et al.* construisent la fonction

$$f(x,y) = \mathbf{1}_{c_t(x,y)<1-\epsilon} + \mathbf{1}_{c_t(x,y)\geq 1-\epsilon}\mathbf{1}_{|\mathbf{v}(x,y)|<T_V}$$

qu'ils régularisent et qu'ils injectent dans les contours actifs géodésiques à la place de la traditionnelle fonction g. Mais pour une telle fonction, l'avantage des snakes géodésiques sur les snakes géométriques est bien moindre. Ici la fonction est une fonction indicatrice lissée, donc lorsqu'elle est nulle ou quasiment nulle (cela dépend du type de lissage que l'on effectue), on a $\dfrac{\partial u}{\partial t} \approx 0$, et donc le

snake n'évolue plus. L'utilisation des contours actifs est même discutable, tant le résultat final dépend du masque trouvé préalablement, les contours actifs servant ici principalement à rendre connexes et à convexifier des zones d'un même objet déconnectées entre elles. En revanche, afin d'affiner le résultat et de rapprocher encore le snake des contours réels des objets, un traitement *a posteriori* basé sur le niveau de gris nécessite d'utiliser les contours actifs géodésiques (*cf.* [KWSR]).

3.4 Un nouveau modèle de segmentation d'objets en mouvement par le flot optique

La dernière méthode de segmentation d'objets en mouvement que nous venons de présenter utilise un flot optique calculé par la méthode des tenseurs. Nous avons comparé visuellement le flot optique donné sur la séquence qui nous intéresse (***cf.* figure 3.24**), pour cela on effectue un seuillage au niveau qui nous paraît le meilleur. Dans [BW], Brox et Weickert avaient comparé pour plusieurs méthodes de flot optique l'erreur commise en calculant le flot optique sur des séquences dont on connaît le flot théorique et avaient conclu que l'erreur la plus faible était obtenue pour la méthode de Weickert et Schnörr. Nos expériences confirment également que cette méthode donne un flot visuellement très correct, et notamment remplit bien les objets en mouvement sans laisser de "trous" notables. En **figure 3.24**, on montre successivement le résultat (l'image sur laquelle le flot optique est calculé est montrée en **figure 3.23**) pour Horn et Schunck avec un faible coefficient ($\alpha = 4$, choisir un coefficient élevé régulariserait immanquablement les contours des objets), Weickert et Schnörr ($\alpha = 750$, les autres paramètres sont usuels et conseillés par Weickert), Weickert et Schnörr avec une procédure de multirésolution ($J = 3$ est la résolution la plus grossière) et la méthode des tenseurs de Bigün améliorée par Weickert que nous venons de présenter dans la section précédente. Signalons la détection sur une zone de l'image en bas à gauche : cette zone n'est pas en mouvement mais un changement d'illumination a lieu sur plusieurs images successives, d'où la détection de cette zone sur le résultat des flots obtenus par la méthode de Horn et Schunck et par celle des tenseurs. Toutefois, ce changement d'illumination n'a lieu que sur quelques images, ce qui n'est pas significatif sur toute la durée de la séquence, c'est ce qui fait que la zone n'est pas détectée comme une zone de mouvement sur le résultat du flot optique obtenu par la méthode de Weickert et Schnörr. Le lissage temporel propre à cette méthode donne un résultat visuellement plus "propre".

3.4.1 Le modèle variationnel

Dans la méthode de segmentation des objets en mouvement de Weickert que nous avons présenté en section précédente, la fonction g est une fonction indicatrice régularisée. La segmentation en elle-même a peu d'effets sur le résultat donné

FIG. 3.23: Une image extraite d'une séquence. Les différents objets en mouvement sont les voitures.

par le seuillage de la norme du flot, si ce n'est la convexification des objets en mouvement. Il nous a donc paru plus naturel d'utiliser une formulation basée région que contours.
Idéalement, nous voudrions que le flot optique soit plus grand en norme qu'un certain seuil à l'intérieur de la région et plus petit que ce seuil à l'extérieur. Compte tenu de la nature particulière du flot optique (mélange des informations de niveaux de gris de plusieurs images), nous allons rajouter également un terme pour attirer le contour vers les bords des objets. La façon la plus naturelle de procéder est de minimiser le long de la frontière de la région la fonction $g(\frac{|\nabla G_\sigma \star I|}{\mu})$ où $g(r) = \frac{1}{1+r^2}$. Malheureusement, cela ne suffit pas et nous verrons que tel quel, les contours obtenus sont trop irréguliers. Pour cette raison, nous avons ajouté un terme pénalisant la longueur dans la fonctionnelle. Pour ne pas alourdir les notations, nous noterons $g(\frac{|\nabla G_\sigma \star I(\mathbf{x})|}{\mu})$ simplement $g_I(\mathbf{x})$. Finalement, la fonctionnelle à minimiser est

$$E(\Omega) = \int_\Omega \alpha \, d\mathbf{x} + \int_{D\setminus\Omega} |\mathbf{v}| \, d\mathbf{x} + \lambda \int_{\partial\Omega} g_I(\mathbf{x}) \, d\mathcal{H}^1(\mathbf{x}) + \nu \int_{\partial\Omega} d\mathcal{H}^1(\mathbf{x})$$

ou bien

$$E_\epsilon(u) = \int_D \alpha H_\epsilon(u) \, d\mathbf{x} + \int_D |\mathbf{v}|(1 - H_\epsilon(u)) \, d\mathbf{x} + \int_D (\lambda g_I + \nu) \, |\nabla u| \delta_\epsilon(u) \, d\mathbf{x}$$

si on représente Ω comme l'ensemble $\{u > 0\}$. En utilisant la technique des dérivées de domaine ou la technique des fonctions de Heaviside, on obtient

$$dE(\Omega; V) = \int_{\partial\Omega} (\alpha - |\mathbf{v}| + (\lambda g_I + \nu)\kappa + \lambda < \nabla g_I, \mathbf{n} >) < V, \mathbf{n} > \, ds$$

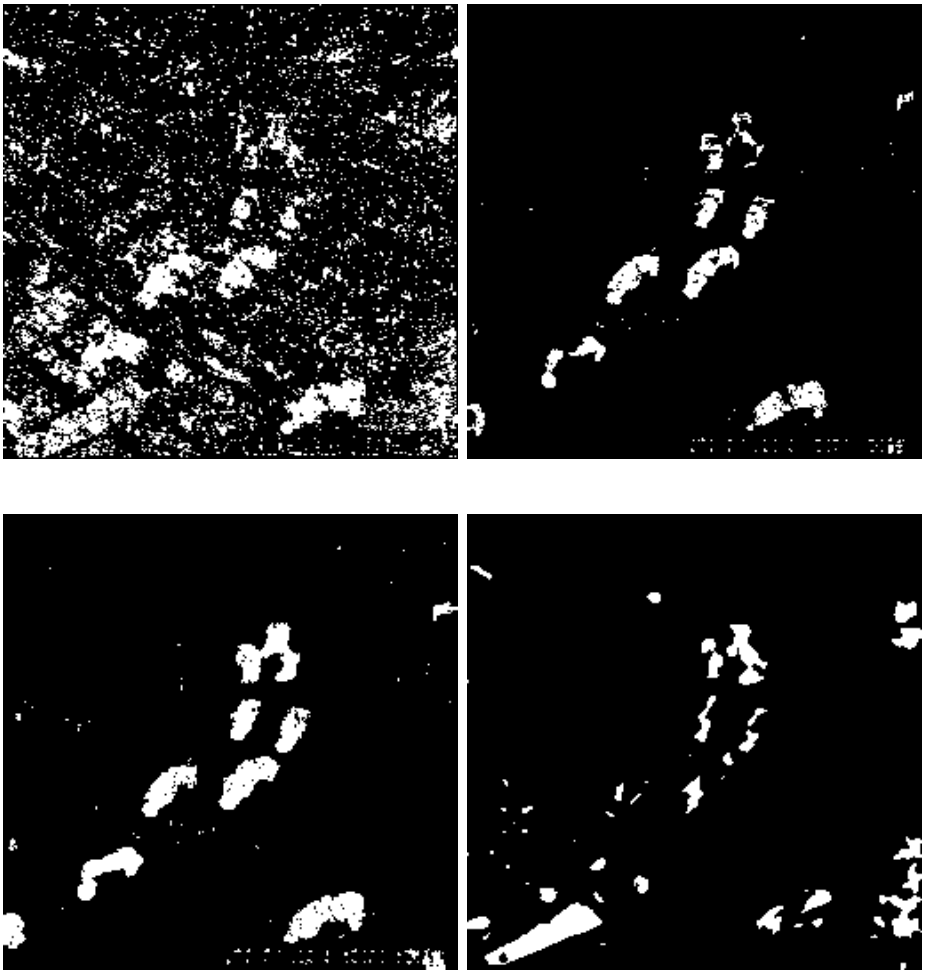


FIG. 3.24: Les masques montrés sont obtenus en seuillant la norme du flot optique sur l'image précédente (*cf.* figure 3.23). De gauche à droite et de haut en bas : Horn et Schunck, Weickert-Schnörr, Weickert-Schnörr avec multirésolution et la méthode des tenseurs de Bigün. Pour les trois premières, le flot est seuillé à $0,7$, pour la dernière à $0,5$ (visuellement le meilleur). Les résultats les plus convaincants sont obtenus pour le modèle de Weickert et Schnörr, que ce soit avec ou sans multirésolution.

ou

$$dE_\epsilon(u;v) = \int_D \left(\alpha - |\mathbf{v}| - (\lambda g_I + \nu)\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) - \lambda < \nabla g_I, \frac{\nabla u}{|\nabla u|} >\right) \delta_\epsilon(u) v \, d\mathbf{x}$$

respectivement.
La minimisation de l'énergie E peut être réalisée par une descente de gradient en choisissant $V = -\left((\lambda g_I(\mathbf{x}) + \nu)\kappa + \lambda < \nabla g_I, \mathbf{n} > +\alpha - |\mathbf{v}|\right)\mathbf{n}$

$$\frac{\partial \Gamma}{\partial t} = -\left((\lambda g_I(\mathbf{x}) + \nu)\kappa + \lambda < \nabla g_I, \mathbf{n} > +\alpha - |\mathbf{v}|\right)\mathbf{n}. \tag{3.6}$$

Si on utilise l'énergie E_ϵ, la minimisation s'effectue en prenant $v = -\left(\alpha - |\mathbf{v}| - (\lambda g_I + \nu)\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) - \lambda < \nabla g_I, \frac{\nabla u}{|\nabla u|} >\right)\delta_\epsilon(u)$, et l'EDP obtenue par descente de gradient est

$$\frac{\partial u}{\partial t} = \delta_\epsilon(u)\left(\mathrm{div}\left((\lambda g_I(\mathbf{x}) + \nu)\frac{\nabla u}{|\nabla u|}\right) + |\mathbf{v}| - \alpha\right). \tag{3.7}$$

3.4.2 Analyse des différents termes de l'énergie et résultats expérimentaux

Nous minimisons donc

$$E(\Omega) = \int_\Omega \alpha \, d\mathbf{x} + \int_{D\setminus\Omega} |\mathbf{v}| \, d\mathbf{x} + \lambda \int_{\partial\Omega} g_I(\mathbf{x}) \, d\mathcal{H}^1(\mathbf{x}) + \nu \int_{\partial\Omega} d\mathcal{H}^1(\mathbf{x}).$$

Analysons les différents termes de cette énergie. Le terme $\int_\Omega \alpha \, d\mathbf{x} + \int_{D\setminus\Omega} |\mathbf{v}| \, d\mathbf{x} + \nu \int_{\partial\Omega} d\mathcal{H}^1(\mathbf{x})$ est très semblable à la fonctionnelle que nous avons vu en section 3.2.1.1 page 41 et vise à trouver un domaine sur lequel $|\mathbf{v}|$ soit assez grand comparé à α, tout en ayant un bord assez régulier.
Quant au terme $\int_{\partial\Omega} g_I(\mathbf{x}) \, d\mathcal{H}^1(\mathbf{x})$, comme nous l'avons déjà dit, il vise à attirer les contours vers les bords des objets. Nous présentons en **figure 3.25**[3] plusieurs résultats qui tendent à prouver que ce terme peut être utile pour bien séparer les objets qui sont regroupés en un seul ensemble si l'on minimise la fonctionnelle sans ce terme. Sur la même figure, nous comparons les résultats obtenus pour la minimisation de la fonctionnelle sans le terme de longueur avec ceux obtenus avec. Les contours sont donc assez irréguliers si l'on n'utilise pas le terme de longueur.
Les **figures 3.28, 3.29** montrent que le résultat du seuillage du flot optique peut être considérablement amélioré si l'on choisit les paramètres de façon appropriée. En effet, on pourra comparer ces figures où le coefficient λ a été choisi de manière plus adéquate ($\lambda = 15$) aux **figures 3.26 et 3.27** ($\lambda = 2$). Le contour est attiré vers les vrais bords des objets en mouvement qui sont correctement séparés les

[3] La séquence est téléchargeable librement sur le site de l'Université de Karlsruhe `i21www.ira.uka.de/image_sequences/`.

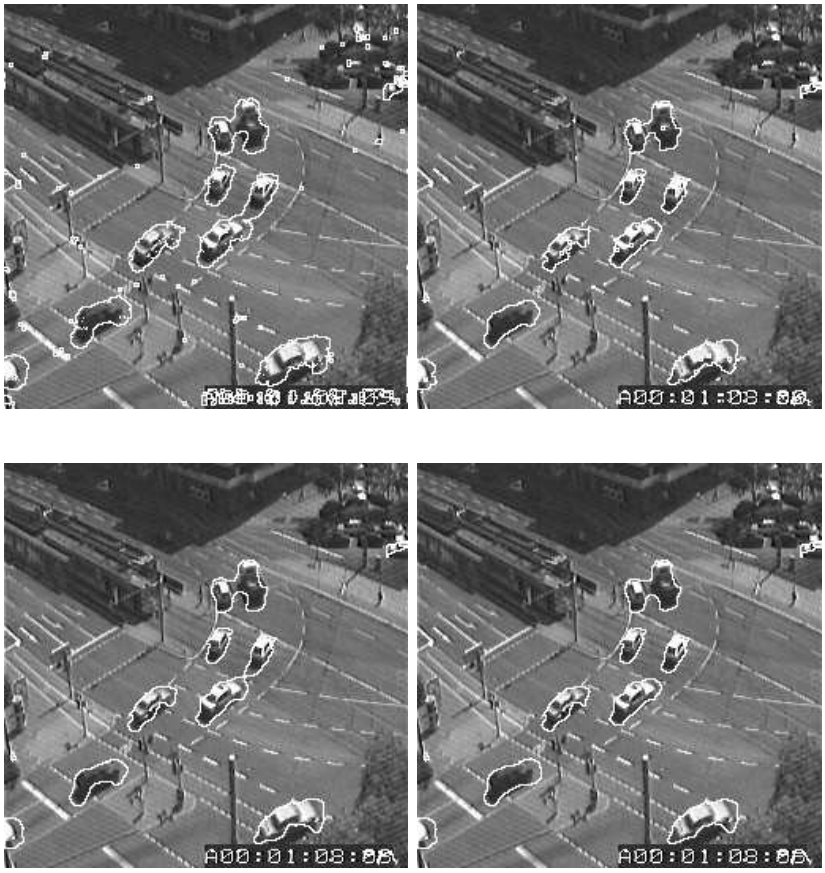


FIG. 3.25: Etude de l'influence de chaque terme dans notre modèle de segmentation avec flot optique. Successivement, le seuillage de la norme du flot à 0.6, le résultat de notre algorithme pour $\lambda = 0$ (sans le terme de longueur) et $\nu = 2$, avec $\lambda = 0,2$ et $\nu = 0$ (sans le terme géodésique), enfin avec $\lambda = 0,2$ et $\nu = 2$. Les autres paramètres sont $\delta t = 5$ pour le pas de temps, $\epsilon = 1$, $\mu = 2$. On voit donc que le terme de longueur est nécessaire pour bien débruiter le masque, mais que le terme de détecteur de bord est utile pour bien séparer les objets en mouvement et pour attirer le contour vers les bords des objets. Il permet également de contrebalancer les effets des zones homogènes (observer la voiture en bas à gauche).

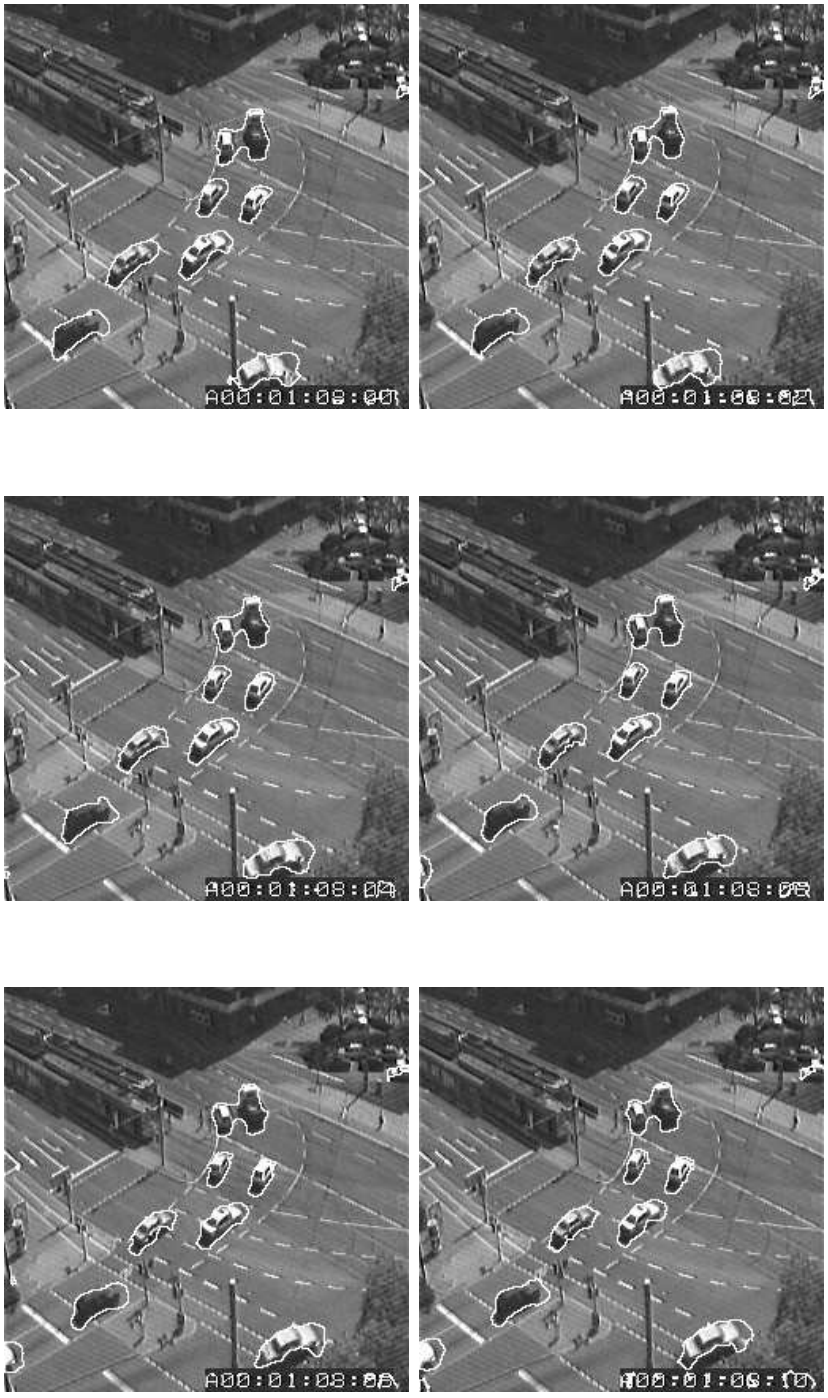


FIG. 3.26: Résultats de notre modèle de segmentation avec flot optique sur les images paires de la séquence (de la 2e à la 12e). Les paramètres sont $\alpha = 0,6$, $\nu = 0,2$, $\lambda = 2$, $\epsilon = 1$ et $\Delta t = 5$ pour le pas de temps.

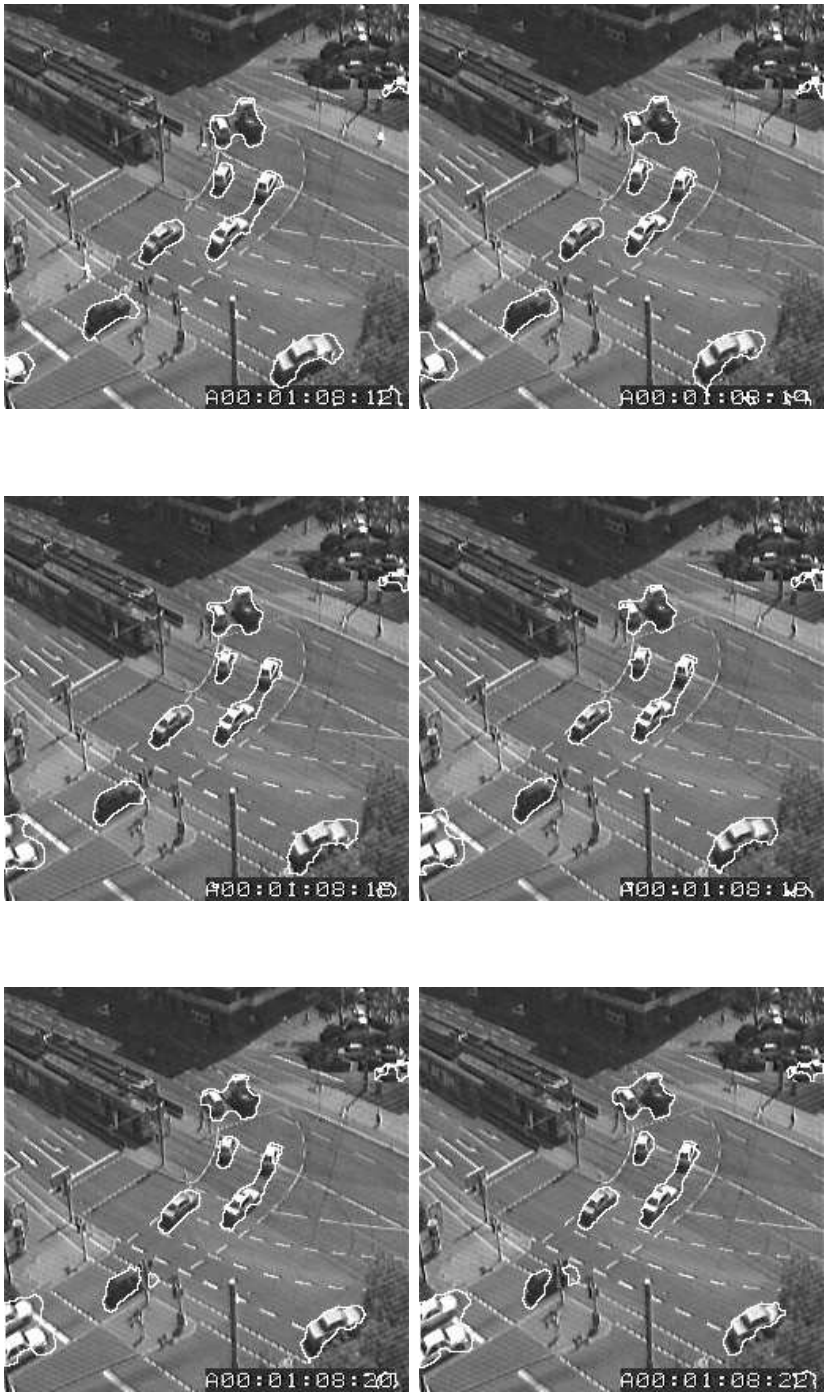


FIG. 3.27: Suite de la figure 3.26. Résultats de notre modèle de segmentation avec flot optique sur les images paires de la séquence (de la 14^{e} à la 24^{e}). Les paramètres sont $\alpha = 0,6$, $\nu = 0,2$, $\lambda = 2$, $\epsilon = 1$ et $\Delta t = 5$ pour le pas de temps.

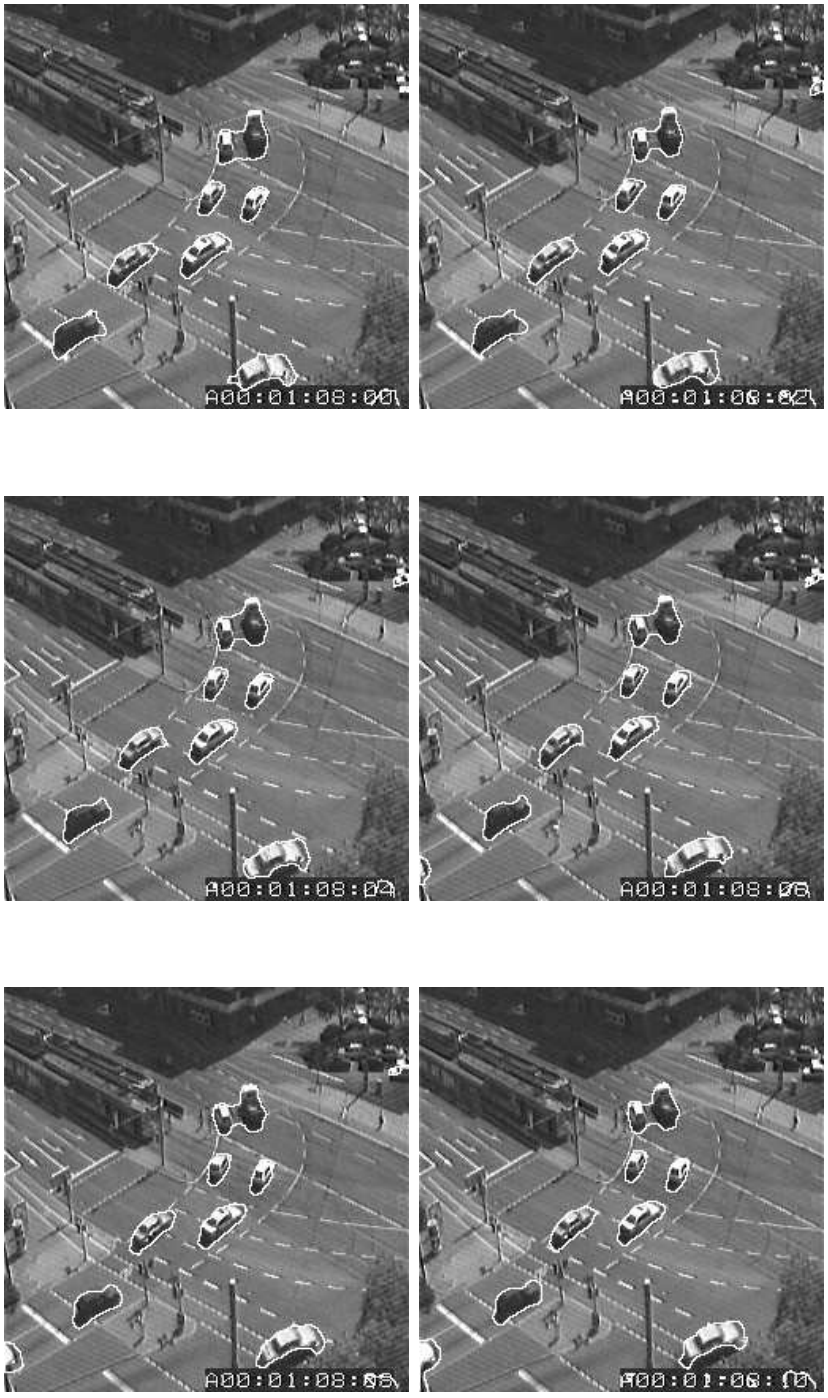


FIG. 3.28: Résultats de notre modèle de segmentation avec flot optique sur les images paires de la séquence (de la 2^{e} à la 12^{e}). Par rapport à la figure 3.26, les paramètres ont été changés en $\nu = 1$ et $\lambda = 15$. Remarquer que l'effet des zones homogènes a disparu (voiture en bas à gauche), le coefficient λ plus élevé qu'en figure 3.26 permettant d'attirer la courbe vers les bords des voitures.

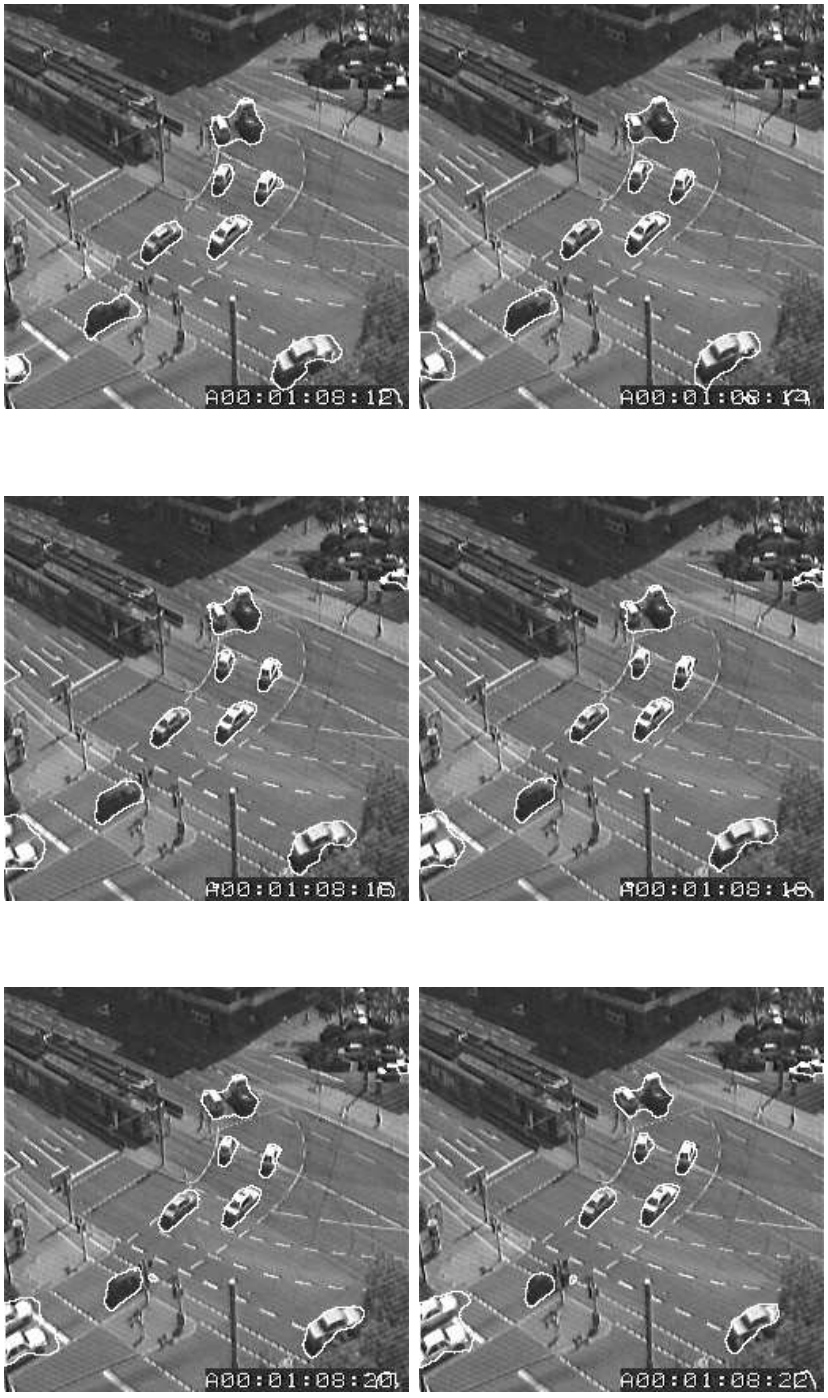


FIG. 3.29: Résultats de notre modèle de segmentation avec flot optique sur les images paires de la séquence (de la 14e à la 24e). Par rapport à la figure 3.27, les paramètres ont été changés en $\nu = 1$ et $\lambda = 15$. Remarquer que l'effet des zones homogènes a disparu (voiture en bas à gauche), le coefficient λ plus élevé qu'en figure 3.27 permettant d'attirer la courbe vers les bords des voitures. On voit aussi que certains objets en mouvement sont parfaitement séparés grâce à un choix judicieux de λ, ce qui n'est pas le cas sur la figure 3.27.

uns des autres saufs quelques-uns qui sont regroupés parce que trop proches. La séquence se prête toutefois bien à l'application de notre algorithme, puisqu'en général les objets se détachent assez bien sur un fond homogène.

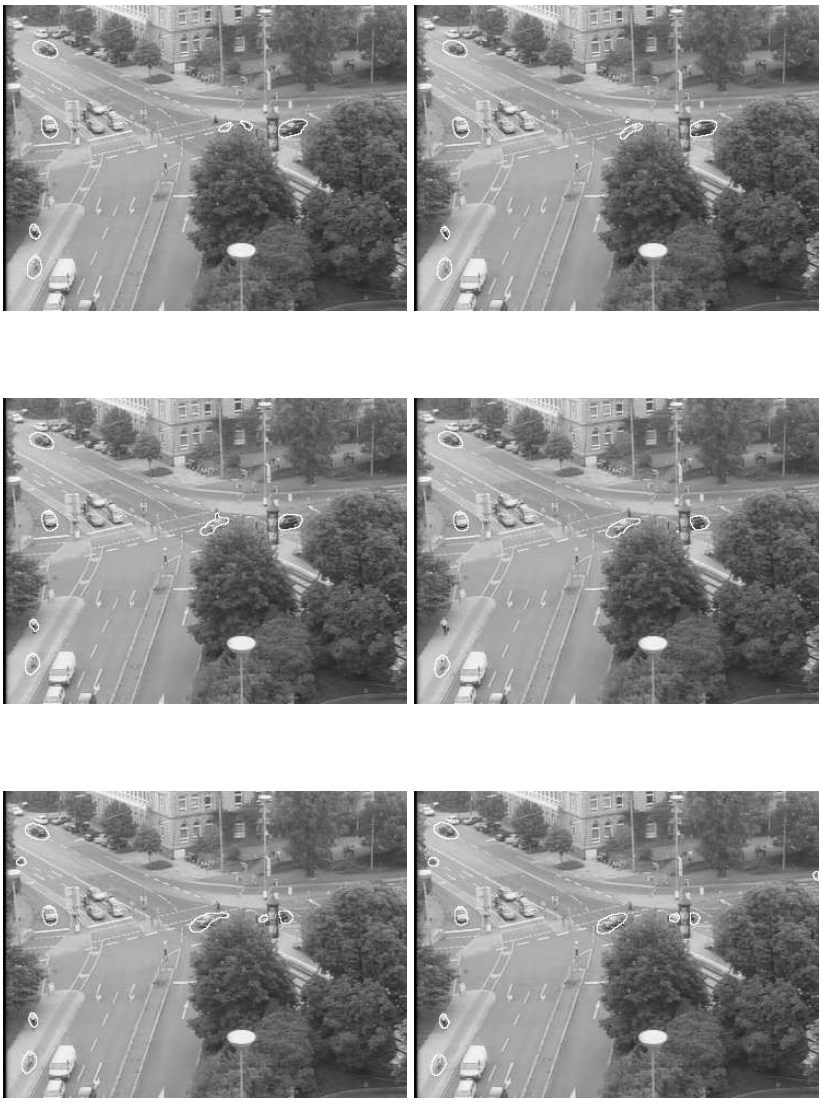


FIG. 3.30: Résultats de notre modèle de segmentation avec flot optique sur les images paires de la séquence (de la 2^{e} à la 12^{e}). Les paramètres sont $\alpha = 0,2$, $\nu = 0,2$, $\lambda = 2$, $\epsilon = 1$ et $\Delta t = 5$ pour le pas de temps.

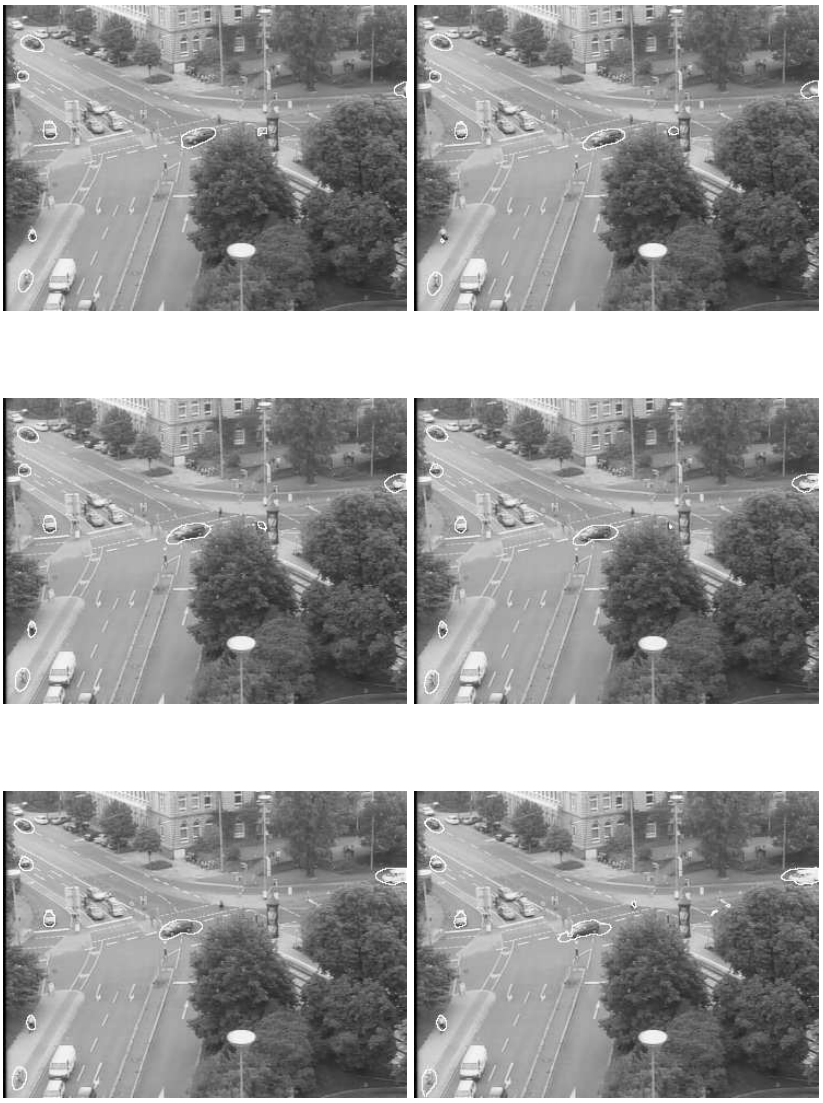


FIG. 3.31: Résultats de notre modèle de segmentation avec flot optique sur les images paires de la séquence (de la 14^{e} à la 24^{e}). Les paramètres sont $\alpha = 0,2$, $\nu = 0,2$, $\lambda = 2$, $\epsilon = 1$ et $\Delta t = 5$ pour le pas de temps.

Chapitre 4

Second cas : la caméra est mobile

Dans ce chapitre, on considère le cas où une caméra mobile filme une scène où un ou plusieurs objets sont en mouvement. Comme on cherche à s'inspirer des méthodes utilisées en segmentation vidéo dans le cas d'une caméra fixe, on a besoin de calculer puis de soustraire le déplacement induit par le mouvement de caméra. Pour cela, on va utiliser une méthode introduite par Odobez et Bouthemy. Cette méthode consiste à trouver un déplacement paramétrique qui réalise le minimum de l'erreur quadratique entre l'image et l'image recalée. La première section de ce chapitre est donc consacrée à l'exposition de cette méthode. Dans tout ce chapitre, elle est utilisée pour recaler les séquences d'images que l'on traite.
Les deux sections suivantes sont dédiées à des méthodes de segmentation par comparaison entre l'image et sa précédente ou entre l'image et un fond donné (pour calculer le fond, on recale la séquence). Comme en section 3.2, on minimise la version discrète des énergies par l'algorithme ICM. On montrera toutefois que l'utilisation de la minimisation de la fonctionnelle continue par EDP reste avantageuse à certains points de vue, notamment à cause d'effets de pixelisation observés sur les résultats de l'ICM. Enfin, dans la dernière section, on utilise le modèle basé sur le flot optique (présenté en section 3.4) sur les séquences recalées.

4.1 Calcul du mouvement de caméra

Dans [OB], Odobez et Bouthemy ont proposé un modèle pour calculer un mouvement paramétrique 2D induit par le déplacement de la caméra. On note $\mathbf{v}_G(\mathbf{x})$ ce déplacement au point $\mathbf{x}$ entre l'image I_t et l'image I_{t+1}. Le meilleur recalage est alors choisi comme le minimum de l'énergie

$$E_{BO}(\mathbf{v}_G, \zeta) = \sum_{\mathbf{x}\in D} \rho\left(I_{t+1}(\mathbf{x} + \mathbf{v}_G(\mathbf{x})) - I_t(\mathbf{x}) + \zeta\right)^2 .$$

L'autre variable ζ est une variable qui tient compte d'un éventuel changement global d'illumination entre les deux images, ρ est une des fonctions ayant de bonnes propriétés pour le lissage au voisinage des discontinuités, par exemple $\rho(s) = \sqrt{\epsilon + s^2}$, $\rho(s) = \frac{s^2}{1+s^2}$.
L'algorithme de minimisation est implémenté avec un procédé de multirésolution : on calcule le minimum $\mathbf{v}^{j+1}$ à la résolution $j+1$, on suréchantillonne ce minimum pour revenir à la résolution j et on cherche les perturbations δ^j suffisamment petites autour de cette valeur $T^j_{j+1}v^{j+1}$ (T^j_{j+1} est l'opérateur de suréchantillonnage) telles que l'on puisse linéariser le terme

$$I^j_{t+1}(\mathbf{x} + T^j_{j+1}v^{j+1} + \delta^j) - I^j_t(\mathbf{x})$$

en

$$r(\delta^j)(\mathbf{x}) = I^j_{t+1}(\mathbf{x} + T^j_{j+1}v^{j+1}) - I^j_t(\mathbf{x}) + < \nabla I^j_{t+1}(\mathbf{x} + T^j_{j+1}v^{j+1}), \delta^j > .$$

Ainsi l'inconnue pour chaque résolution j est δ^j et l'énergie à minimiser

$$E^j(\delta^j, \zeta^j) = \sum_{\mathbf{x}\in D} \rho\left(r(\delta^j)(\mathbf{x})\right).$$

Odobez et Bouthemy utilisent alors le théorème semi-quadratique énoncé dans [CBFAB] et l'algorithme de minimisations alternées suivant.

1. Pour $k \in \mathbb{N}$, à partir de l'erreur r^k commise à l'itération k, on calcule
$$w^k = \frac{\rho'(r^k(\delta^j)(\mathbf{x}))}{r^k(\delta^j)(\mathbf{x})},$$
2. on minimise par rapport à δ_j
$$\sum_{\mathbf{x}\in D} w^k \left(r^{k+1}(\delta^j)(\mathbf{x})\right)^2$$
3. jusqu'à convergence.

On suppose $\mathbf{v}$ paramétrique de degré au plus égal à deux. On a donc en général

$$\mathbf{v}(x,y) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} q_1 & q_2 & q_3 \\ q_4 & q_5 & q_6 \end{pmatrix} \begin{pmatrix} x^2 \\ xy \\ y^2 \end{pmatrix}.$$

Le mouvement paramétrique comprend donc au plus 12 paramètres, l'intérêt des termes d'ordre deux apparaissant lorsque la déformation subie par l'image est différente d'une similitude du plan mais comporte une partie projective. Dans ce cas, le mouvement paramétrique à 12 paramètres n'en est encore qu'une approximation, mais une approximation suffisante en pratique. Par ailleurs, il a été démontré par F. Dibos que la déformation de l'image dûe au déplacement de la caméra peut être exprimé en fonctions de 6 paramètres seulement ([Di]) au lieu des 8 du groupe projectif, il y a donc une redondance d'information entre les différents paramètres de la méthode que nous venons d'exposer.
Nous tenons à remercier l'équipe "Vista" de l'Irisa/Inria Rennes pour l'utilisation du logiciel Motion2D.

4.2 Soustraction du mouvement induit par le déplacement de la caméra et segmentation

4.2.1 Segmentation par comparaison entre deux images successives

Une fois calculé le mouvement de caméra par la méthode exposée dans la section précédente, on construit à partir de chaque image de la séquence l'image suivante prédite à partir de l'image et de la connaissance des paramètres du mouvement entre l'image et sa suivante. On a donc

$$\tilde{I}_{t+1}(\mathbf{x}) = I_t(\mathbf{x} - \mathbf{v})$$

et on compare $\tilde{I}_{t+1}$ et I_t.
Nous allons introduire la même fonctionnelle que dans [ABJB3]

$$E_{ABJB}(\Omega) = \int_{\Omega} \alpha \, d\mathbf{x} + \int_{D \setminus \Omega} |\tilde{I}(t+1, \mathbf{x}) - I(t, \mathbf{x})| \, d\mathbf{x} + \lambda \int_{\partial\Omega} d\mathcal{H}^1. \qquad (4.1)$$

On cherche donc le domaine assez régulier (on minimise entre autre la longueur de Hausdorff de sa frontière) pour lequel la différence entre l'image prédite $\tilde{I}_{t+1}$ et l'image précédente I_t est assez faible à l'extérieur du domaine et plus grande qu'un certain seuil α à l'intérieur.

Tout comme dans la section 3.2.1.1, nous avons implémenté la minimisation de (4.1) en utilisant l'ICM. Comme nous l'avions expliqué, il y a peu de différences en pratique entre une implémentation par minimisation locale de l'énergie discrète et une implémentation par une minimisation plus globale (comme à la façon de Chan/Vese dans [CV1]), surtout en choisissant comme condition initiale le résultat du seuillage de $|\tilde{I}_{t+1}(\mathbf{x}) - I_{t+1}(\mathbf{x})|$ par α, car cette initialisation est *a priori* proche de la solution du problème de minimisation de (4.1) (c'est la solution de la minimisation avec $\lambda = 0$).
Sur la séquence montrée sur la **figure 4.1**[1], on cherche à détecter le mouvement d'un joueur de tennis en faisant l'hypothèse qu'une fois le mouvement de caméra enlevé, la différence entre l'image prédite et l'image réelle soit d'une amplitude suffisamment grande sur la zone du joueur. Cette hypothèse n'est que partiellement vérifiée, nous avons constaté en effet que certaines régions sur le joueur étaient trop homogènes pour que la différence des niveaux de gris des images soit significative. Nous avons donc choisi un seuil α qui nous donne un résultat visuellement correct sur le joueur mais ayant comme inconvénient que beaucoup d'autres régions de l'image sont détectées (on peut donc assimiler ces régions à du bruit, puisqu'elles sont détectées alors qu'elles ne devraient pas l'être, voir

[1] Nous tenons à remercier Michel Barlaud et son équipe de l'I3S pour nous avoir fourni la séquence "Edberg" ainsi que d'autres séquences.

figure 4.3.
Pour les résultats, nous avons calculé le mouvement de caméra sur une durée de 90 images successives et avons recalé le film par rapport à sa première image. Nous avons donc une nouvelle séquence $\tilde{I}$ à caméra fixe (*cf.* **figure 4.2**). Ensuite, on calcule les différences entre deux images consécutives $\tilde{I}_t$ et $\tilde{I}_t$. Nous n'avons donc pas exactement suivi le modèle d'Aubert *et al.* mais cette formulation est équivalente puisque dans les deux cas, on effectue une compensation du mouvement. Le modèle correspondant est alors

$$E_{ABJB}(\Omega) = \int_{\Omega} \alpha \, d\mathbf{x} + \int_{D\setminus\Omega} |\tilde{I}(t+1,\mathbf{x}) - \tilde{I}(t,\mathbf{x})| \, d\mathbf{x} + \mu \int_{\partial\Omega} d\mathcal{H}^1. \qquad (4.2)$$

Comme on l'a dit, la difficulté de ce modèle tient dans le traitement des zones homogènes en mouvement. Dans le cas de notre joueur de tennis, son tea-shirt est une zone uniformément blanche de taille assez grande, et donc la différence de niveau de gris est assez faible d'une image à sa suivante, il est donc difficile de ne pas avoir des "trous" dans notre segmentation. Afin d'éviter ce problème, on est conduit à diminuer le seuil α, et donc le résultat de la segmentation est très bruité (*cf.* **figure 4.3**). La solution que nous avons proposée est d'effectuer un filtre de grain sur le résultat pour enlever toutes les composantes connexes des ensembles de niveaux de taille inférieure à un certain seuil (*cf.* **figure 4.4**).

4.2.2 Segmentation par comparaison entre l'image et le fond

Ainsi que nous l'avons indiqué dans la section précédente, à partir de la connaissance du mouvement de la caméra, on peut comparer chaque image avec l'image prédite à partir d'une autre image. Nous avons choisi de faire ceci pour toutes les images : on construit une séquence à caméra fixe où chaque image de la séquence est recalée sur l'image courante I_k et on calcule le fond B de cette séquence. On peut alors appliquer les algorithmes que nous avons vu en section 3.2.1.1 sur la comparaison entre une image et le fond. On rappelle que l'on cherche le domaine Ω qui minimise l'énergie

$$E(\Omega) = \int_{\Omega} \alpha \, d\mathbf{x} + \int_{D\setminus\Omega} |B - I_k|(\mathbf{x}) \, d\mathbf{x} + \lambda \int_{\partial\Omega} d\mathcal{H}^1. \qquad (4.3)$$

L'énergie discrète correspondante est alors

$$E(\phi) = \sum_{i,j} (\alpha - |B - I_k|(i,j)) H(\phi(i,j)) + \lambda \sum_{i,j} \underbrace{\sqrt{(H(\phi(i+1,j)) - H(\phi(i,j)))^2 + (H(\phi(i,j+1)) - H(\phi(i,j)))^2}}_{\triangleq l(i,j)}.$$

Nous avons réutilisé l'ICM que nous avons présenté en section 3.2.1.1 dans le cas présent. Cela constitue à notre avis une alternative raisonnable à la méthode de

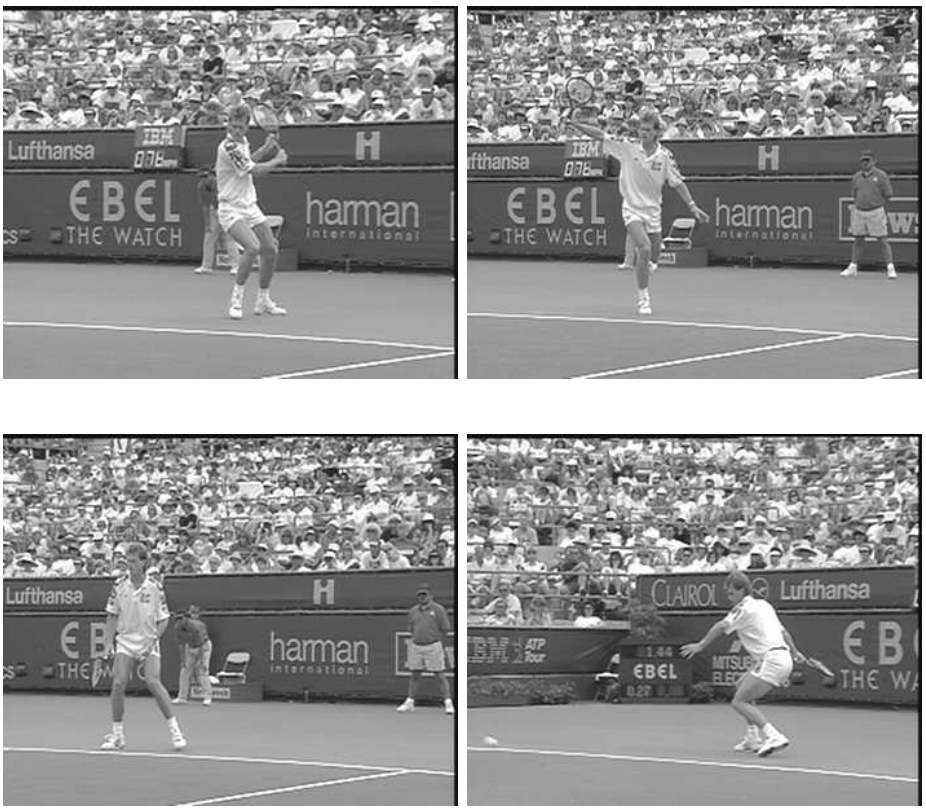


FIG. 4.1: Quelques images de la séquence “Edberg”.

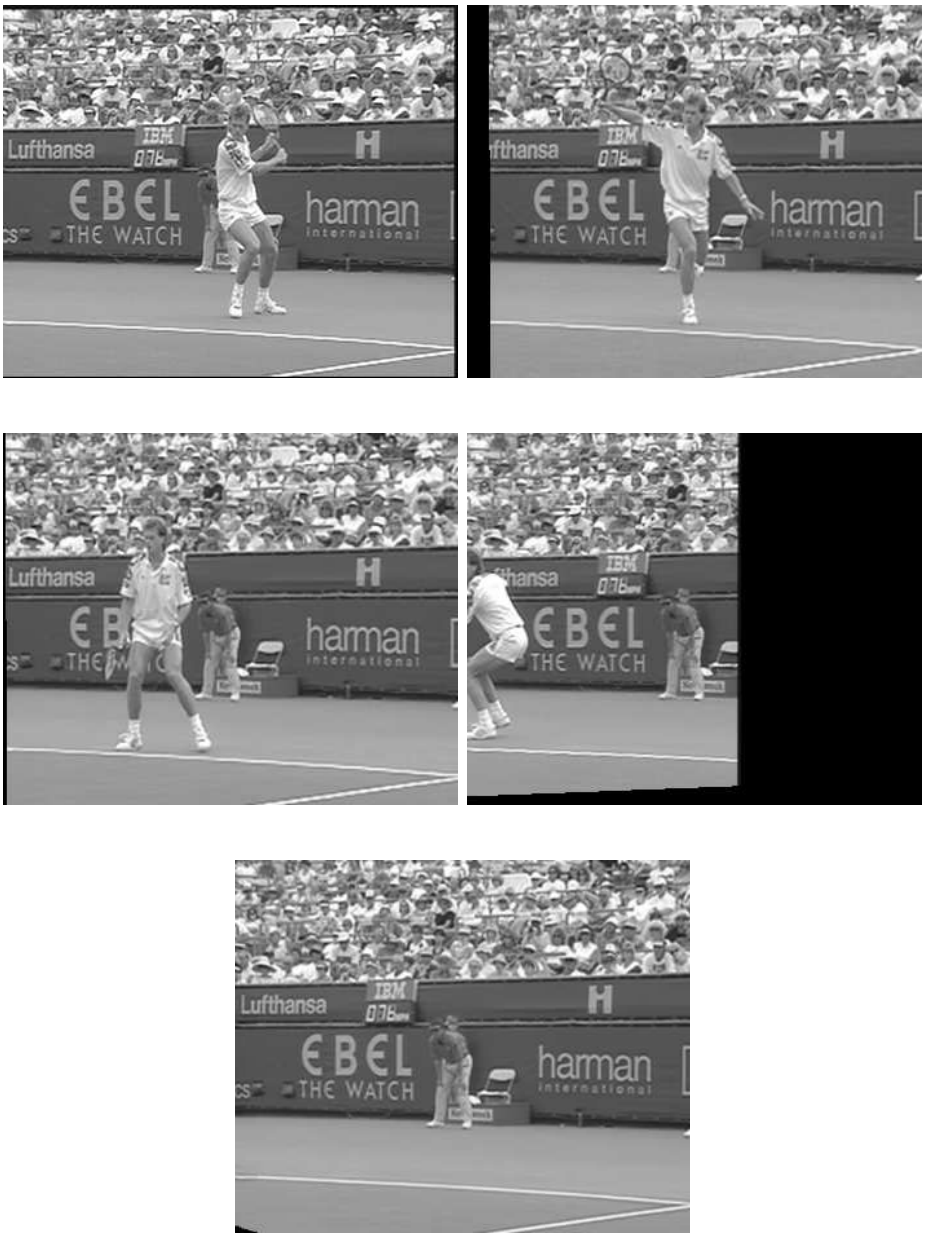


FIG. 4.2: Quelques images de la séquence recalée, puis le fond calculé sur cette séquence.

contours actifs par EDP qui est utilisée par Barlaud *et al.* ([ABJB2]), en effet, l'énergie à minimiser n'est que la version discrète de l'énergie de 4.3 et l'ICM a l'avantage d'être beaucoup plus rapide qu'une méthode par implémentation d'une EDP. De plus, dans [BP], M. Barlaud et F. Precioso notaient que les algorithmes par EDP sont en règle générale plutôt lents, pour cette raison ils avaient proposé d'appliquer le même algorithme mais à un contour actif paramétrique, en l'occurence une courbe spline, afin d'accélérer la vitesse de l'algorithme.
Ici nous avons calculé le fond par un filtre médian temporel, c'est-à-dire qu'en chaque pixel, le niveau de gris du fond est la médiane de l'ensemble des niveaux de gris des images en ce pixel, ce qui s'exprime par

$$B(\mathbf{x}) = \underset{k\in\{1,\dots,n\}}{\text{med}}(I_{t_k}(\mathbf{x})) = \begin{cases} \frac{\tilde{I}_{t_p}(\mathbf{x})+\tilde{I}_{t_{p+1}}(\mathbf{x})}{2} & \text{si } n = 2p \\ \tilde{I}_{t_{p+1}}(\mathbf{x}) & \text{si } n = 2p+1 \end{cases}$$

où $\tilde{I}_{t_k}(\mathbf{x})$ est la suite ordonnée des $I_{t_k}(\mathbf{x})$

$$\tilde{I}_{t_1}(\mathbf{x}) \leq \tilde{I}_{t_2}(\mathbf{x}) \leq \dots \leq \tilde{I}_{t_n}(\mathbf{x}).$$

Rappelons également que le choix d'un filtre médian temporel a une interprétation variationnelle, puisque la médiane temporelle des images est la solution du problème variationnel

$$\inf_B \sum_k \int_D |B - I_{t_k}|(\mathbf{x})\, d\mathbf{x}.$$

Comparé à la moyenne des images, qui correspond à la solution du problème variationnel

$$\inf_B \sum_k \int_D |B - I_{t_k}|^2(\mathbf{x})\, d\mathbf{x},$$

la médiane à la propriété bien connue d'être moins sensible aux valeurs extrêmes des intensités que la moyenne.
Les résultats obtenus sont évidemment meilleurs que ceux obtenus à partir du modèle exposé dans la section précédente, puisque nous disposons de la donnée du fond qui est une donnée plus fiable. Quelques résultats sont présentés en **figure 4.5** sur plusieurs images consécutives (comparer avec la **figure 4.4**). Le fond est montré en **figure 4.2**. Pour avoir un résultat encore plus correct, nous avons utilisé la procédure de décroissance des paramètres introduite par Jehan-Besson ([ABJB1]) :

1. on effectue l'algorithme pour une valeur α_1 du paramètre α,
2. le résultat obtenu sert d'initialisation pour réappliquer l'algorithme avec une valeur plus faible $\alpha_2 < \alpha_1$,
3. on itère ce procédé, l'initialisation étant toujours le résultat précédent, et $\alpha_n < \alpha_{n+1}$.

Pour nos expériences, nous avons choisi trois valeurs de α ($\alpha_1 = 30$, $\alpha_2 = 20$ et $\alpha_3 = 10$). Afin d'obtenir un meilleur résultat, nous avons encore ajouté un post-processing de filtre de grain avec un seuil (50), mais cette dernière étape pourrait être optionnelle, le résultat étant déjà de bonne qualité sans ce traitement.
A titre de comparaison, nous montrons en **figure 4.6** le résultat obtenu en utilisant les mêmes paramètres dans la minimisation de la fonctionnelle continue par EDP. On constate que le contour est plus régulier (il y a moins d'effets de blocs) que sur le résultat de l'ICM, où l'irrégularité du contour est visuellement un peu gênante. Là plus que dans les résultats sur les séquences de voitures, on voit les différences entre un algorithme local comme l'ICM et un algorithme plus global tel que l'EDP. Comme nous l'avons vu en section 3.2.1.3, on peut alors combiner les deux algorithmes. Cela permet de réduire le temps de calcul de l'algorithme par EDP car la solution donnée par l'application de l'ICM est plus proche du résultat final. La plupart du temps, les résultats obtenus par la combinaison des deux algorithmes sont quasiment identiques à ceux obtenus par l'EDP seule (*cf.* **figures 4.6 et 4.7**) mais les temps de calcul sont radicalement différents. Si l'on considère par exemple le temps de calcul de l'EDP seule dans notre cas où nous utilisons trois valeurs de α successives, les temps sont respectivement (sur la 43^{e} image de la séquence) de 14 s, 38 s et 114 s pour chacune des trois valeurs. Lorsque l'on applique préalablement l'ICM, les temps de calcul de l'EDP sont alors de 40 s, 26 s et 56 s : on réduit quasiment de moitié le dernier temps de calcul qui est le plus long des trois. La réduction du temps de calcul est fluctuante selon les images, mais est toujours en faveur de la combinaison des deux algorithmes. Sur certaines images, on constate même un meilleur résultat avec la combinaison qu'en utilisant seulement l'algorithme par EDP (*cf.* **figures 4.8 et 4.9**).
En comparant **les figures 4.3 et 4.5**, on voit deux avantages à comparer les images plutôt avec le fond : d'une part, les paramètres donnant un résultat visuellement correct ont des valeurs assez naturelles à trouver, d'autre part, même si le joueur n'a pas un déplacement très important sur les images que l'on montre, la comparaison avec le fond suffit à le discriminer assez bien, alors que la comparaison avec les images prédites à partir de l'image précédente échoue (même en choisissant des paramètres assez faibles comme dans la section précédente) à cause du déplacement du joueur qui est trop faible et à cause des zones homogènes en niveau de gris sur le joueur.

En conclusion, on peut dresser un bilan des avantages et des inconvénients des deux méthodes que nous venons de présenter dans les deux sections que nous venons de voir.
– Pour la méthode de comparaison d'une image avec l'image prédite à partir de l'image précédente, on peut obtenir de bons résultats mais à condition de choisir un paramètre de seuillage plutôt faible pour bien pouvoir détecter les

objets en mouvement (ici le joueur) et si le mouvement des objets n'est pas trop faible. Le résultat ainsi obtenu est ensuite débruité par l'application d'un filtre de grain.
- Pour la méthode de comparaison d'une image avec un fond, on obtient de bons résultats sans choisir les paramètres de la façon la plus judicieuse possible. Le choix de plusieurs paramètres de seuillage, des valeurs élevées vers les valeurs les plus faibles, permet de débruiter le résultat puis de l'affiner sur les objets en mouvement.
- Tous les algorithmes que nous avons appliqué dans ce chapitre (calcul de fond par filtre médian temporel, segmentation par ICM) sont particulièrement rapides (de l'ordre de quelques secondes), leur temps de calcul est donc très faible. De plus, si l'on souhaite bénéficier des avantages de l'algorithme par EDP, le temps de calcul de celui-ci peut être réduit par une application préalable de l'ICM, comme nous l'avons montré en section 3.2.1.3.
- Se pose la question de la stabilité des paramètres tout au long de la séquence. Peut-on choisir les mêmes paramètres pour toutes les images de la séquence et obtenir de bons résultats ? Selon toute vraisemblance oui, à condition qu'il n'y ait pas de variations de luminosité importante, ce qui rendrait de toute façon le calcul de fond difficile voire impossible.

4.3 Segmentation par utilisation du flot optique résiduel

Dans cette section, on reprend le modèle que nous avons exposé dans le chapitre 3 section 3.4. On recale les images de la séquence en utilisant la méthode de recalage de Bouthemy pour recaler les ensembles d'images qui peuvent l'être. On a donc une séquence à caméra fixe à traiter. On peut donc utiliser la méthode de segmentation par flot optique que nous avons vu en section 3.4 : on calcule le flot optique sur la séquence recalée par la méthode de Weickert et Schnörr, puis on minimise

$$E(\Omega) = \int_{\Omega} \alpha \, d\mathbf{x} + \int_{D\setminus\Omega} |\mathbf{v}| \, d\mathbf{x} + \lambda \int_{\partial\Omega} g(\mathbf{x}) \, d\mathcal{H}^1(\mathbf{x}) + \nu \int_{\partial\Omega} d\mathcal{H}^1(\mathbf{x}).$$

Pour nos expériences, nous avons utilisé un calcul de flot optique multirésolution. En effet, dès lors que les objets en mouvement sont de taille importante, il convient que les zones homogènes en mouvement soient bien détectées comme du mouvement, ce qui n'est pas nécessairement évident avec un flot optique calculé sans compensation des grands mouvements comme cela est fait dans un calcul multirésolution.

Nous présentons nos expériences sur les **figures 4.10 et 4.14** sur deux séquences recalées. A titre de comparaison, la **figure 4.13** montre le résultat que l'on obtiendrait si on seuillait la norme du flot optique au niveau α pour la séquence montrée en **figure 4.14**. Cela montre que notre résultat dépend à un certain

FIG. 4.3: Résultats du modèle basé sur la différence entre deux images successives sur plusieurs images consécutives de la séquence "Edberg" ($\alpha = 10$, $\lambda = 50$, 5 balayages de l'ICM). Le résultat est assez bruité et comporte des "trous" car certaines zones (le T-shirt) sur le joueur sont trop homogènes en niveau de gris.

FIG. 4.4: Résultats du modèle basé sur la différence entre deux images successives avec un post-processing de filtre de grain ($\alpha = 10$, $\lambda = 50$, 5 balayages de l'ICM). Le résultat de la figure précédente est visuellement amélioré par le filtre de grain.

FIG. 4.5: Résultats de l'implémentation par ICM sur le modèle basé sur la différence entre une image et le fond ($\alpha_1 = 30$, $\alpha_2 = 20$, $\alpha_3 = 10$, $\lambda = 75$, 5 balayages de l'ICM). Le résultat est évidemment bien meilleur qu'en utilisant simplement la différence avec l'image précédente (comparer avec la figure 4.3).

FIG. 4.6: Résultats de l'implémentation par EDP sur le modèle basé sur la différence entre une image et le fond ($\alpha_1 = 30$, $\alpha_2 = 20$, $\alpha_3 = 10$, $\lambda = 75$, $\delta t = 5$ pour le pas de temps). Le résultat est meilleur, grâce notamment à la disparition d'effets de blocs (*aliasing*) du contour.

FIG. 4.7: Résultats de l'application successive de l'ICM et de l'EDP sur le modèle basé sur la différence entre une image et le fond ($\alpha_1 = 30$, $\alpha_2 = 20$, $\alpha_3 = 10$, $\lambda = 75$, $\delta t = 5$ pour le pas de temps). Le résultat est quasiment identique au résultat de l'EDP seule : l'avantage de la combinaison des deux algorithmes réside principalement dans la réduction du temps de calcul.

FIG. 4.8: Modèle basé sur la différence entre une image et le fond ($\alpha_1 = 30$, $\alpha_2 = 20$, $\alpha_3 = 10$, $\lambda = 75$, $\delta t = 5$ pour le pas de temps). Cas où le résultat de l'EDP seule (deuxième image) est amélioré par l'application préalable de l'ICM (troisième image). La première image montre le résultat de l'ICM seul.

FIG. 4.9: Modèle basé sur la différence entre une image et le fond ($\alpha_1 = 30$, $\alpha_2 = 20$, $\alpha_3 = 10$, $\lambda = 75$, $\delta t = 5$ pour le pas de temps). Cas où le résultat de l'EDP seule (deuxième image) est amélioré par l'application préalable de l'ICM (troisième image). La première image montre le résultat de l'ICM seul.

point de la qualité du flot optique obtenu, ce que nous avions déjà souligné dans la section 3.4 du chapitre 3. On voit également que le résultat n'est pas parfait pour plusieurs raisons : le contour présente des artefacts qui proviennent très souvent des phénomènes d'occultation (une région n'a pas de région correspondante dans l'image suivante ou l'image précédente), ou bien seule une partie de l'objet est en mouvement, ce qui induit une détection sur cette partie et non sur tout l'objet (*cf.* **figures 4.10 et 4.11**). Par rapport au modèle de comparaison des images avec le fond, notre modèle nous donne une information moins riche : alors que le premier modèle nous dit que les objets détectés par l'algorithme se détachent sur le fond, notre modèle nous dit que les objets détectés ont un mouvement assez grand. Cela explique donc les différences de résultats particulièrement flagrants lorsque l'on compare les **figures 4.12 et 4.11**. En revanche, l'algorithme donne des résultats raisonnables sur la séquence[2] montrée en **figure 4.14** car les mouvements sont beaucoup plus rigides, et l'application de notre modèle peut davantage être recommandée dans un tel cas.

[2]librement téléchargeable à partir du site `http://www.cipr.rpi.edu/resource/sequences/sif.html`

FIG. 4.10: Résultats de notre méthode de segmentation par flot optique pour les paramètres suivants : $\alpha = 1$, $\lambda = 5$, $\nu = 1$. Les mouvements non rigides ne favorisent pas l'utilisation de cette méthode. L'emploi d'un flot optique obtenu avec multirésolution permet toutefois d'obtenir un résultat certes approximatif, mais qui peut se combiner avec d'autres algorithmes.

FIG. 4.11: Résultats de notre méthode de segmentation par flot optique pour les paramètres suivants : $\alpha = 1$, $\lambda = 25$, $\nu = 5$. Les mouvements non rigides ne favorisent pas l'utilisation de cette méthode. A comparer avec la figure précédente (les paramètres sont plus élevés, la courbe est donc plus proche des bords des objets).

FIG. 4.12: Résultats obtenus pour le modèle de segmentation par comparaison entre l'image et le fond en appliquant l'ICM et l'EDP ($\alpha_1 = 30$, $\alpha_2 = 20$, $\alpha_3 = 10$, $\lambda = 75$, $\delta t = 5$ pour le pas de temps). A comparer avec la figure précédente.

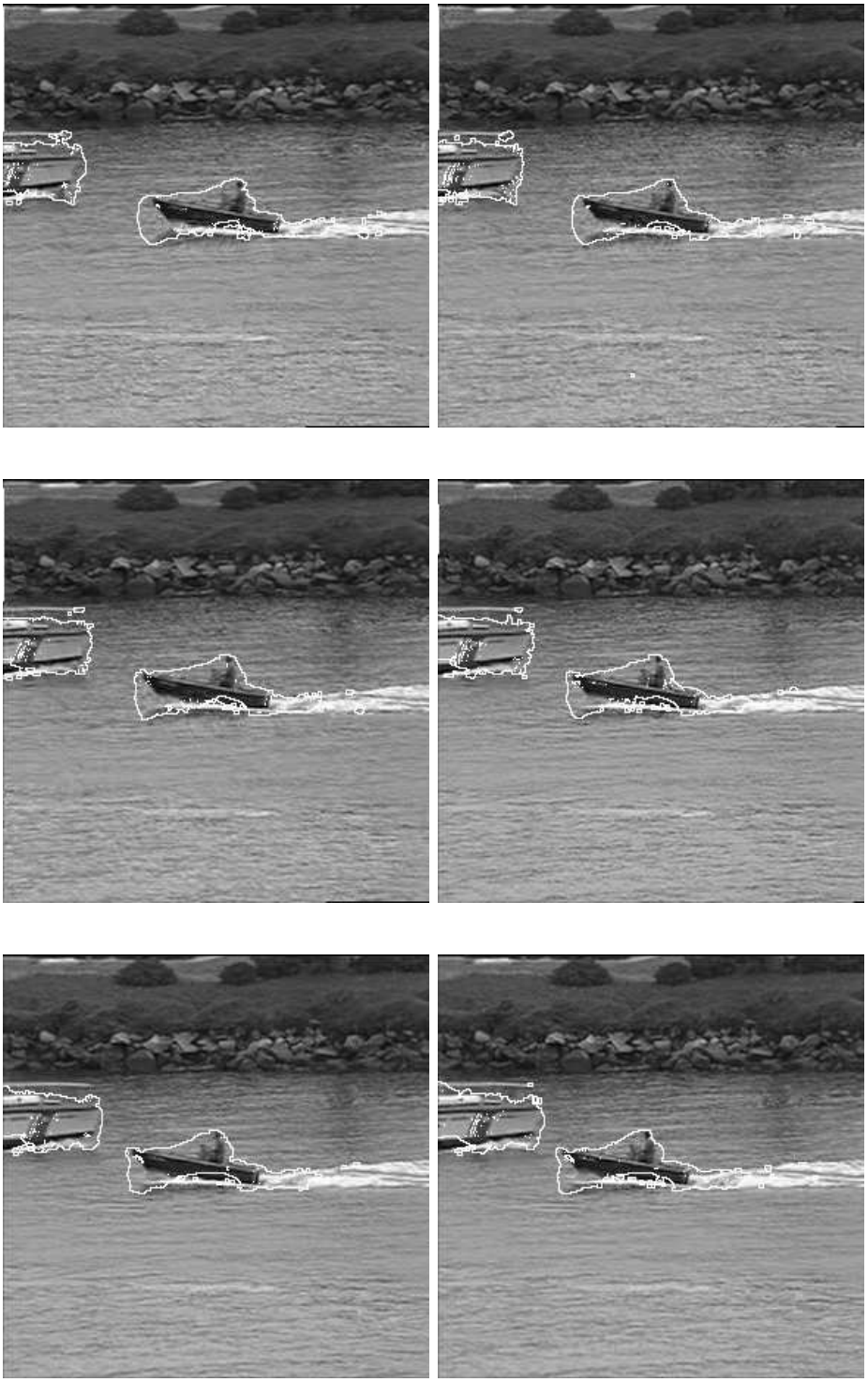


FIG. 4.13: Résultat du seuillage de la norme du flot à $1,5$.

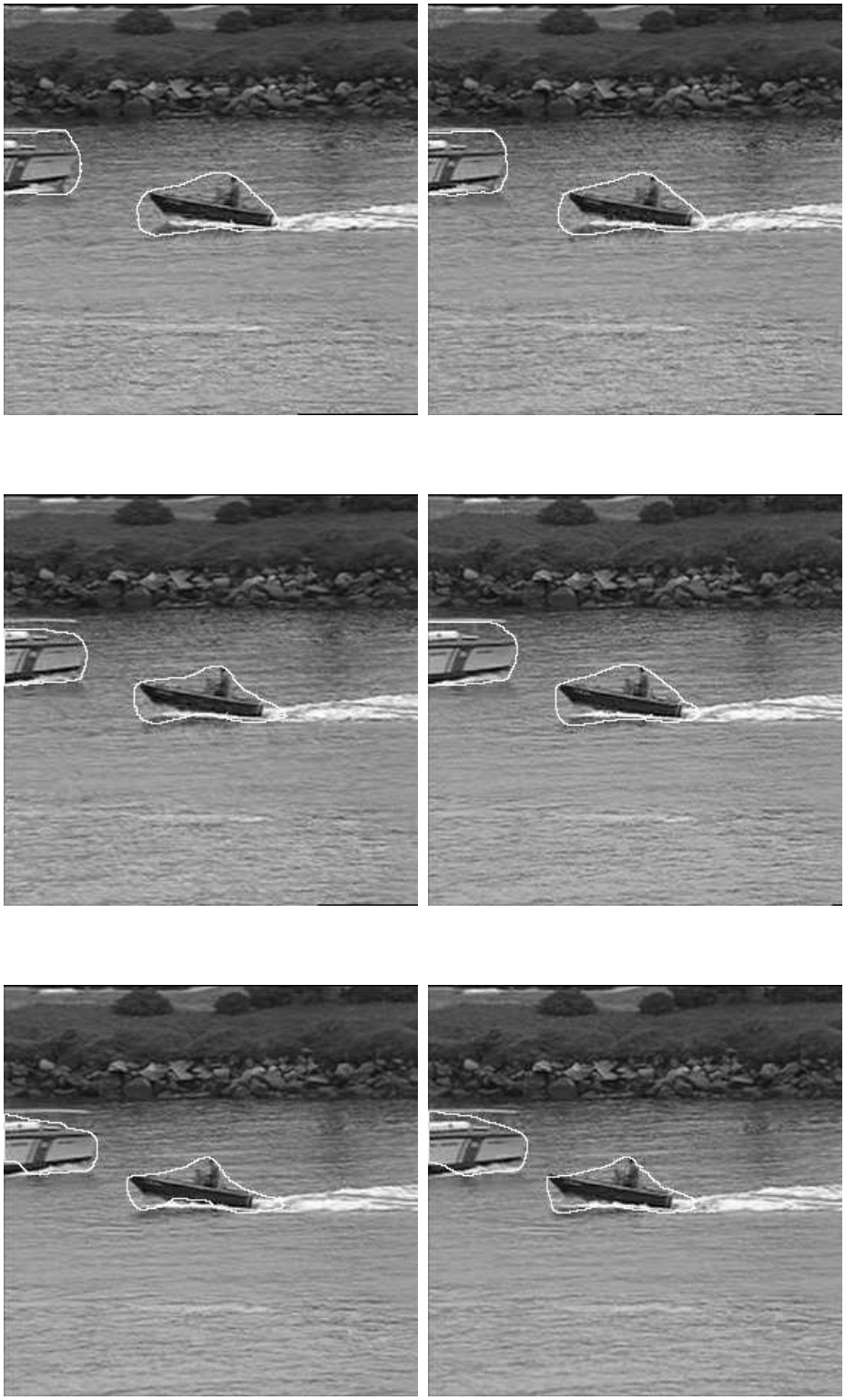


FIG. 4.14: Résultats de notre méthode de segmentation par flot optique pour les paramètres suivants : $\alpha = 1,5$, $\lambda = 20$, $\nu = 5$. Cette méthode peut donc être particulièrement indiquée pour des séquences d'images où les mouvements sont assez rigides.

Chapitre 5

Analyse vidéo en temps réel

Ce chapitre est consacré à l'étude de deux problèmes de vidéo surveillance qui nous ont été soumis. Il s'agit de l'estimation de la densité d'une foule et de la détection de comportements anormaux de personnes dans l'environnement du métro parisien. Les travaux exposés dans ce chapitre ont été réalisés dans le cadre d'un contrat entre le CEREMADE et la RATP. Nos travaux ont été validés par l'entreprise VISIOWAVE qui fournit à la RATP le matériel ainsi que les applications de traitement d'image sous forme de DLL. Dans ce contrat, le CEREMADE devait étudier et fournir des solutions aux deux problèmes que sont le comptage de personnes sur un quai ainsi que la détection de mouvements anormaux de personnes dans les metros.
On se donne une séquence d'images, la scène filmée étant typiquement un quai, un couloir ou un escalator du métro parisien. On veut déterminer quasiment en temps réel, si à un moment donné, des personnes ont un comportement *anormal*. Par comportement *anormal*, on entend que la projection des vitesses sur le plan image (le *flot optique*) donne des directions différentes de directions prédéfinies par un opérateur.
Dans la première section, on présente le principe d'une méthode actuellement utilisée pour la détection de changements dans une vidéo, et nous expliquons pourquoi son utilisation est limitée pour le problème de détection des comportements anormaux. Ensuite, on étudiera une classe de méthodes de calcul de flot fréquemment appelées méthodes locales en exposant leurs avantages. N'ayant pas été exposées dans les chapitres précédents, ces méthodes seront introduites dans la deuxième partie du chapitre. Leur utilisation est motivée par le contexte de la vidéosurveillance qui nécessite des algorithmes rapides. Contrairement à la méthode de détection de changements, elles peuvent être envisagées pour la détection de comportements anormaux dans un environnement impliquant un certain nombre de difficultés (quai de métro...). Enfin, on exposera les résultats concernant les directions du mouvement et la détection effective de comportement anormal.

5.1 Détermination des objets en mouvement par détection de changement

Une méthode développée à l'origine par Aach dans [AKP] et utilisée par Ziliani *et al.* dans [MCZM] consiste à écrire que la différence de deux images successives I_h et I_{h+1} est la réalisation d'une variable aléatoire gaussienne aux points où il n'y a pas de mouvement. On a donc, en notant H_0 l'hypothèse "il n'y a pas de mouvement au point (x,y)", la densité de probabilité conditionnelle

$$p_{I_{h+1}-I_h|H_0} = \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{(I_{h+1}-I_h)^2(x,y)}{2\sigma^2}).$$

Les auteurs de [MCZM] indiquent que la variance σ^2 peut être choisi comme le double de la variance du bruit gaussien s'ajoutant aux images, que l'on peut connaître connaissant le type de caméra que l'on manipule.
A partir de là, la détection des objets en mouvement repose sur un test statistique. On calcule la somme des carrés des $(\frac{I_{h+1}-I_h}{\sigma})^2$ dans une fenêtre 5×5 centrée sur le pixel courant

$$\Delta_{(x,y)} = \sum_{(x',y')\in\{x-2,\ldots,x+2\}\times\{y-2,\ldots,y+2\}} \left(\frac{I_{h+1}(x',y') - I_h(x',y')}{\sigma}\right)^2.$$

D'après la loi de probabilité écrite ci-dessus, sous l'hypothèse d'indépendance des

$$\left(\frac{I_{h+1}(x',y') - I_h(x',y')}{\sigma}\right)^2,$$

$\Delta_{(x,y)}$ suit une loi du χ^2 à 25 degrés de liberté puisqu'il y a 25 pixels dans la fenêtre. Dans le cas d'une fenêtre de taille $N \times N$, $\Delta_{(x,y)}$ suit une loi du χ^2 à N^2 degrés de liberté. La démonstration repose sur le théorème classique :

Théorème 7 *Soient n variables aléatoires $X_1, ..., X_n$ indépendantes identiquement distribuées suivant une loi gaussienne centrée réduite : $X_1 \sim \mathcal{N}(0,1)$. Alors la variable aléatoire définie par $X = X_1^2 + ... + X_n^2$ suit une loi gamma de paramètres $\frac{n}{2}$ et $\frac{1}{2}$*

$$X \sim \gamma(\frac{n}{2}, \frac{1}{2})$$

et cette loi est appelée loi du χ^2 à n degrés de libertés.

On rappelle la définition de la loi gamma :

Définition 1 *On dit qu'une variable aléatoire X suit une loi gamma $\gamma(a,\lambda)$ si sa densité de probabilité est donnée par*

$$p_X(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} \exp(-\lambda x) \mathbf{1}_{x>0}$$

Ainsi, on choisit l'état "le pixel appartient à une zone qui a changée" en fonction d'une valeur de référence α et de la valeur $p(\Delta(x,y) > t|H_0)$ où t est un seuil à fixer par l'utilisateur. α est choisi de sorte que l'on ait accès au seuil correspondant t_α dans les tables du χ^2.
Ziliani *et al.* ont utilisé une stratégie multirésolution pour rendre les résultats plus robustes. L'image I à l'échelle l au pixel (x,y) est ainsi obtenue en pondérant les niveaux de gris des pixels de I à l'échelle $l-1$ sur un voisinage 5×5

$$I_l(x,y) = \sum_{k_x=-2}^{2} \sum_{k_y=-2}^{2} w(k_x,k_y) \cdot I_{l-1}(2x+k_x, 2y+k_y)$$

où w est un noyau à choisir.
A partir de là, les auteurs expliquent que le masque (fond en noir et changements en blanc) est sous-échantillonné sur des blocs de 16×16 pixels, puis que chaque bloc de la zone de changement (en blanc sur le masque) est gardé tel quel si tous ces voisins appartiennent encore à cette zone. Le résultat ainsi trouvé s'avère très intéressant pour la compression, ainsi que le montrent les auteurs dans [MCZM] en comparant le PSNR d'une séquence compressée à partir du masque trouvé successivement par le test statistique brut (décision en fonction de $p(\Delta(x,y) > t|H_0)$), le test statistique et la multirésolution, le masque trouvé en examinant les blocs, un masque "idéal" trouvé à la main ou par une méthode sophistiquée. Nous avons implémenté cette méthode de manière brute sans le procédé de multirésolution et d'affinage expliqué ci-dessus. On voit sur la **figure 5.1** que des effets de blocs apparaissent, ce qui est naturel puisque la fenêtre que nous utilisons est de taille fixe (en général 5×5 dans nos expériences).
Une autre approche utilisant la même modélisation est dûe à Paragios et Tziritas ([PT]). En effet, ils supposent toujours que $I_{h+1} - I_h|H_0$ et $I_{h+1} - I_h|H_1$ suivent des lois gaussiennes et donc (on note H_1 l'hypothèse contraire à H_0 : il y a mouvement au point (x,y))

$$p(I_{h+1} - I_h = d) = P_0 p(I_{h+1} - I_h|H_0) + P_1 p(I_{h+1} - I_h|H_1).$$

A partir de là, ils estiment les variances des deux gaussiennes par maximum de vraisemblance puis utilisent les champs de Markov pour labeliser les pixels (statique ou mobile). Cet algorithme a été utilisé par nos partenaires de VISIO-WAVE pour la détection de mouvements anormaux. La méthode suivie consiste en plusieurs étapes

1. Indiquer manuellement une ligne de séparation sur l'image. On cherche à détecter combien de personnes traversent cette ligne et s'ils la traversent dans un sens indésirable (comportement anormal).
2. Détecter les zones en mouvement sur chaque image.
3. Apparier chaque zone de l'image I_h avec une zone dans l'image I_{h+1} selon un critère simple (distance entre les barycentres).

FIG. 5.1: Détection des changements pour une fenêtre 5×5, successivement pour $\sigma = 15$ et $\sigma = 19$.

4. Détecter les objets traversant la ligne et le sens du mouvement à travers la ligne.

Toutefois cette méthode est limitée compte tenu des objectifs, elle est notamment quasi inutilisable si on cherche à détecter les mouvements anormaux dans une séquence de foule : le choix d'emplacement de la ligne de séparation est purement arbitraire et des zones englobant plusieurs personnes ayant des mouvements de direction différentes risquent d'être détectées. Une mesure locale du mouvement comme le flot optique semble plus indiquée pour notre problème.

5.2 Méthodes locales de flot optique

Dans cette section, nous présentons les méthodes locales de calcul du flot optique. Elles reposent sur l'hypothèse que le flot optique est constant sur un voisinage de taille fixe de chaque pixel. Il s'agit donc d'une hypothèse très différente de celle d'une régularité *globale* en un sens mathématique bien défini, hypothèse utilisée dans les méthodes différentielles. Les méthodes locales n'ont pas été présentées dans les chapitres sur la segmentation d'objets en mouvement car elles donnent en général une estimation moins bonne de la norme du flot optique, du moins dans leur version originale. Le travail récent de Kornprobst *et al.* [LKM] pallie à ce défaut en combinant la version originale des méthodes locales avec des stratégies multiéchelles et multigrilles. En revanche, les méthodes locales sont réputées pour donner une bonne estimation de la direction du flot optique et c'est précisément la direction qui va nous intéresser dans l'étude de la détection de comportements anormaux. D'autre part les méthodes différentielles sont très

coûteuses en temps de calcul et sont quasiment inutilisables dans le contexte de la vidéosurveillance qui nécessite des algorithmes rapides. Les méthodes locales sont au contraire très peu coûteuses en temps de calcul puisque le flot optique est calculé en résolvant un système linéaire 2×2 ou 3×3 en chaque point.

5.2.1 La formulation de Lucas et Kanade

Toutes les méthodes locales se basent sur l'hypothèse que le flot optique est constant sur un certain voisinage de chaque pixel. Historiquement, Lucas et Kanade furent les premiers à formuler cette hypothèse, ce qui les conduisit à écrire que le flot optique V devait minimiser la fonctionnelle

$$E_{X_0}(V) = \int_{D(X_0,\rho)} |\nabla I \cdot V + I_t|^2(X)\, dX$$

en tout point X_0 ($D(X_0, \rho)$ désigne le disque de centre X_0 de rayon ρ). La minimisation de cette fonctionnelle donne

$$\left(\int_{D(X_0,\rho)} \nabla I (\nabla I)^T(X)\, dX \right) V = - \left(\int_{D(X_0,\rho)} I_t \nabla I(X)\, dX \right). \tag{5.1}$$

V est donc la solution de ce système linéaire au sens classique si la matrice

$$S = \int_{D(X_0,\rho)} \nabla I (\nabla I)^T(X)\, dX$$

est inversible ; dans le cas contraire, si cette matrice est non nulle, on choisit V comme la solution du système aux moindres carrés

$$V = \operatorname{argmin} \left\| \left(\int_{D(X_0,\rho)} \nabla I (\nabla I)^T(X)\, dX \right) V + \left(\int_{D(X_0,\rho)} I_t \nabla I(X)\, dX \right) \right\|^2$$

où $I(X)$ représente l'intensité de l'image au point X, ∇I son gradient et I_t la dérivée temporelle de l'image (la différence entre deux images consécutives lorsque l'on considère une séquence d'images). En décomposant V sur la base des vecteurs propres de S, on obtient que la solution aux moindres carrés est donnée par

$$V = \frac{b_{max}}{\lambda_{max}} e_{max}$$

où b_{max} est la composante de $-\left(\int_{D(X_0,\rho)} I_t \nabla I(X)\, dX\right)$ sur le vecteur propre e_{max} correspondant à la plus grande valeur propre λ_{max} de S.

5.2.2 Introduction d'un filtrage

Posé tel quel, cette formulation reposant sur une fenêtre $D(X_0, \rho)$, on peut s'attendre à ce que le flot optique présente trop de discontinuités et de sensibilité au bruit. De plus, rien ne garantit que l'on puisse résoudre le système linéaire (5.1), même au sens des moindres carrés. Pour cette raison, on choisit en général de remplacer la fenêtre par un noyau w

$$E_{X_0,w}(V) = \int_\Omega w(X_0 - X)|\nabla I \cdot V + I_t|^2(X)\, dX = w \star |\nabla I \cdot V + I_t|^2$$

où $w \star |\nabla I \cdot V + I_t|^2$ désigne la convolution de $|\nabla I \cdot V + I_t|^2$ par le noyau w. Pour w, on peut prendre simplement une gaussienne, c'est-à-dire

$$w(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{\|X\|^2}{2\sigma^2}),$$

ce qui réduit déjà considérablement le risque que l'on ne puisse pas donner de solutions au système ; on peut aussi suivre C. Bernard qui choisit des ondelettes (*cf.* section suivante). Après introduction du noyau w, le système linéaire à résoudre devient

$$\left(w \star (\nabla I (\nabla I)^T)\right) V = -w \star (I_t \nabla I). \tag{5.2}$$

Cela revient donc à effectuer un préfiltrage du tenseur S et du membre de droite $w \star (I_t \nabla I)$. Concernant les problèmes évoqués ci-dessus, le préfiltrage réduit les effets du bruit et diminue le risque que S soit nulle et donc non inversible.

5.2.3 Utilisation des ondelettes (C. Bernard, [Ber])

Les filtres utilisés dans la méthode de Christophe Bernard sont des ondelettes. Il s'agit d'une frame d'ondelettes complexes (pour des raisons de stabilité) discrètes $(\psi^s_{jkk'})_{s\in S,j,k,k'\in\mathbb{Z}}$ de $L_2(\mathbb{R}^2, \mathbb{C})$, où

$$\psi^s_{jkk'} = 2^j \psi^s(2^j x - k, 2^j y - k').$$

L'indice j représente l'échelle, (k, k') la translation de la fenêtre dans l'image et $S = \{1, 2, 3, 4, 5\}$. $s \in \{1, 2, 3, 4\}$ représente classiquement l'ensemble des orientations des filtres : horizontale, verticale, diagonale et anti-diagonale. $s = 5$ correspond à une cinquième ondelette de moyenne non nulle qui est rajoutée afin de détecter les changements d'illumination.
Classiquement, on utilisera un schéma d'ondelettes dyadiques et séparables pour que le filtrage puisse se faire successivement sur les lignes et les colonnes de l'image. L'équation du flot optique qui en découle est maintenant sur-déterminée :

$$\langle I, \frac{\partial \psi^s_{jkk'}}{\partial x}\rangle v^j_x + \langle I, \frac{\partial \psi^s_{jkk'}}{\partial y}\rangle v^j_y = \frac{\partial}{\partial t}\langle I, \psi^s_{jkk'}\rangle$$

Le système obtenu est donc de la forme $A_j V^j = B_j$ au point $2^{-j}(k, k')$, pour chaque échelle j. La matrice A_j est de taille 5×2 (il y a cinq ondelettes). Classiquement, V^j étant réel, la solution de ce système sur-déterminé est donné par $V^j = (\mathrm{Re}(A_j^T A_j))^{-1} \cdot \mathrm{Re}(A_j^T \cdot B_j)$ où A_j^T est la transposée de la matrice A_j.

5.3 Comparaison avec une méthode différentielle

La méthode alternative que nous avons essayé a été introduite par J. Weickert et C. Schnörr dans [WS]. C'est une des nombreuses méthodes *différentielles* qui existent.
Il s'agit de minimiser la fonctionnelle

$$E(\sigma) = \int_{\Omega\times[0,T]} |\nabla I \cdot \sigma + I_t|^2 \, dx\, dy\, dt + \alpha \int_{\Omega\times[0,T]} \Psi(|\nabla_3 V_1|^2 + |\nabla_3 V_2|^2) \, dx\, dy\, dt$$

où $\lambda > 0$, $\Psi(s^2) = \epsilon s^2 + (1-\epsilon)\lambda^2\sqrt{1 + \frac{s^2}{\lambda^2}}$ et $\nabla_3 = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial t}\right)^T$.
La différence de cette approche par rapport à d'autres approches classiques est l'intégration par rapport au temps qui impose que le flot soit régulier par rapport aux variables spatiales *et* à la variable temporelle.
Cette approche, bien que plus lourde en raison de l'addition de la variable temps, s'est révélée particulièrement efficace notamment pour diminuer les effets dus au bruit tout en conservant les discontinuités du flot.
Une autre différence par rapport à d'autres approches ([ADK1] par exemple) est que l'on intègre $\Psi(|\nabla_3 V_1|^2 + |\nabla_3 V_2|^2)$, ce qui signifie que V_1 et V_2 ont des discontinuités aux mêmes points. On peut penser *a priori* que cette différence de modélisation ne change pas considérablement les résultats sur des images naturelles, mais en fait il se trouve que sur les images sur lesquelles l'algorithme a été testé, la direction semble ne pas bien correspondre à ce à quoi on s'attendrait.

5.4 Expériences, critère d'anormalité

On présente sur les **figures 5.2 et 5.3** une comparaison entre la méthode de Lucas et Kanade avec noyau gaussien et la méthode de Weickert et Schnörr : l'histogramme des directions du flot optique est représenté, et de manière étonnante, ceux donnés par Lucas-Kanade correspondent mieux à l'idée que l'on se fait des directions (les pics correspondant au mouvement de la foule et de la personne à la canne sont bien en $\frac{\pi}{2}$ et $-\frac{\pi}{2}$ alors que pour les histogrammes de Weickert-Schnörr, cela est moins évident).
Enfin, d'un point de vue pratique, si l'on cherche à atteindre le temps réel, il devient irréalisable de le faire en utilisant l'algorithme de Weickert et Schnörr qui nécessite d'être appliqué sur au moins une dizaine d'images afin de suffisamment tenir compte de la variable temporelle, en revanche, le calcul de flot de

Lucas et Kanade filtré est très peu coûteux en temps de calcul puisqu'il consiste à inverser une matrice 2×2 en tout point, le temps de calcul s'élève à $1s$ en moyenne entre deux images consécutives.
Enfin, l'algorithme de Bernard donne des résultats raisonnables mais sans améliorer particulièrement les résultats obtenus par la méthode de Lucas et Kanade. On constate expérimentalement que le choix du noyau de convolution (gaussien ou ondelette) ne change pas considérablement le résultat du flot optique.
Dans cette séquence, des personnes empruntent un escalator, et, à un moment, l'un d'entre eux rebrousse chemin, suivi peu après par les deux autres. On cherche donc dans cette séquence à détecter si un mouvement à contre-sens du mouvement de l'escalator devient non négligeable. Pour cela, on se donne un intervalle de direction privilégié $\mathcal{I}_{\hat{k}} = [\theta_{\hat{k}}, \theta_{\hat{k}+1}[$ dans l'histogramme des directions du flot optique. Pour chaque intervalle $\mathcal{I}_k = [\theta_k, \theta_{k+1}[$, la valeur de l'histogramme est donné par

$$H_k = |\{\mathbf{x}|\, \theta(\mathbf{x}) \in \mathcal{I}_k\}|.$$

Le critère d'anormalité est alors le suivant : si $H_{\hat{k}}$ est supérieur à un certain seuil τ *a priori* à régler *ad. hoc.*, alors on considère que l'on a à faire à une situation anormale. Concrètement, on mesure donc la quantité de pixels pour lesquels la direction se situe dans l'intervalle prédéfini et si cette quantité est suffisamment élevée, on déclenche une alarme. Une fois un τ convenable trouvé, on peut choisir de le garder pour d'autres séquences d'images de la même scène s'il n'y a pas de changement de luminosité ou de phénomène trop gênant pour le calcul du flot optique.
Les **figures 5.4, 5.5, 5.6 et 5.7** montrent l'histogramme des directions pour la méthode de Bernard à différents moments de la descente. On voit que de manière logique le pic dans la direction $-\frac{\pi}{2}$ croît au fur et à mesure que la personne descend. Les résultats sont encourageants, puisque sur 700 images de la séquence de l'escalator (*cf.* **figure 5.4**), on a

- 6 fausses alarmes, autrement dit une alarme est générée alors qu'il n'y a aucun comportement anormal ;
- sur plus de 200 images de la descente de la personne de la **figure 5.8**(c), une alarme est générée, le comportement est bien anormal selon notre modélisation ;
- sur les 20 premières images de la descente, il n'y a pas de comportement anormal signalé.

Nous avons réalisé le même test par la méthode de Lucas et Kanade, et on peut dire que les résultats sont comparables, puisqu'on a

- environ 30 fausses alarmes ;
- sur environ les dix premières images de la descente, il n'y a pas de comportement anormal signalé ;
- sur le reste des images de la descente, le comportement anormal est effectivement détecté.

Il nous semble donc que l'avantage de cette dernière réside dans le faible temps de calcul.

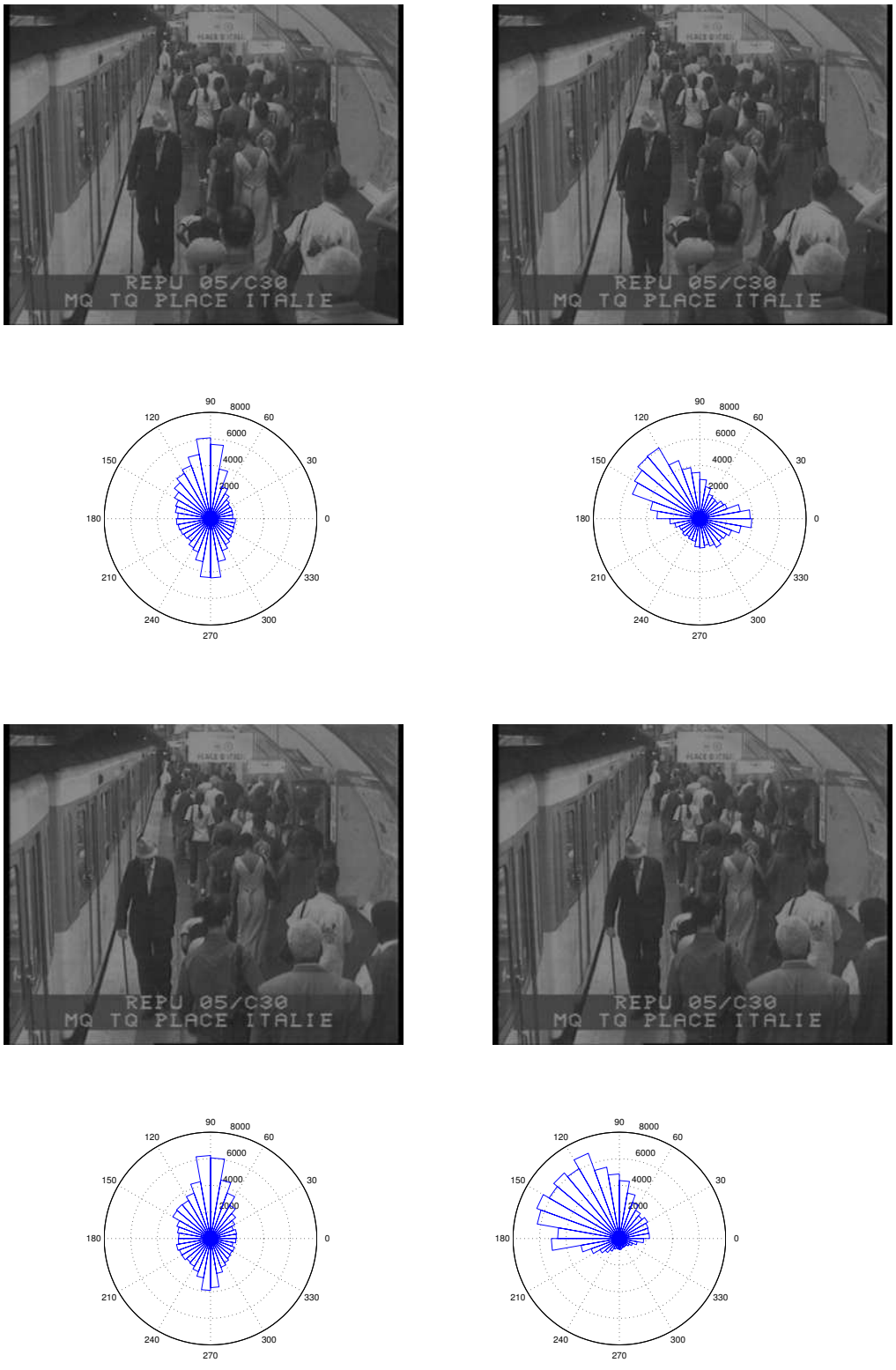


FIG. 5.2: Plusieurs couples d'images sur lesquels le flot optique est calculé ; histogrammes des directions (Lucas-Kanade à gauche, Weickert-Schnörr à droite).

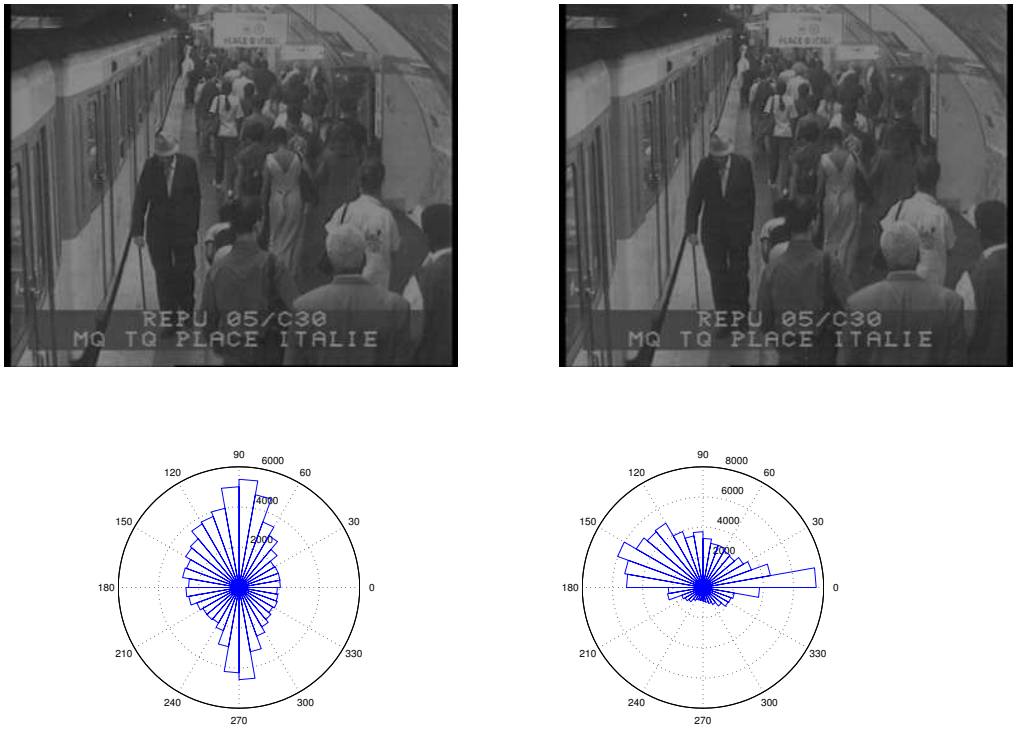


FIG. 5.3: Même commentaire que sur la figure précédente.

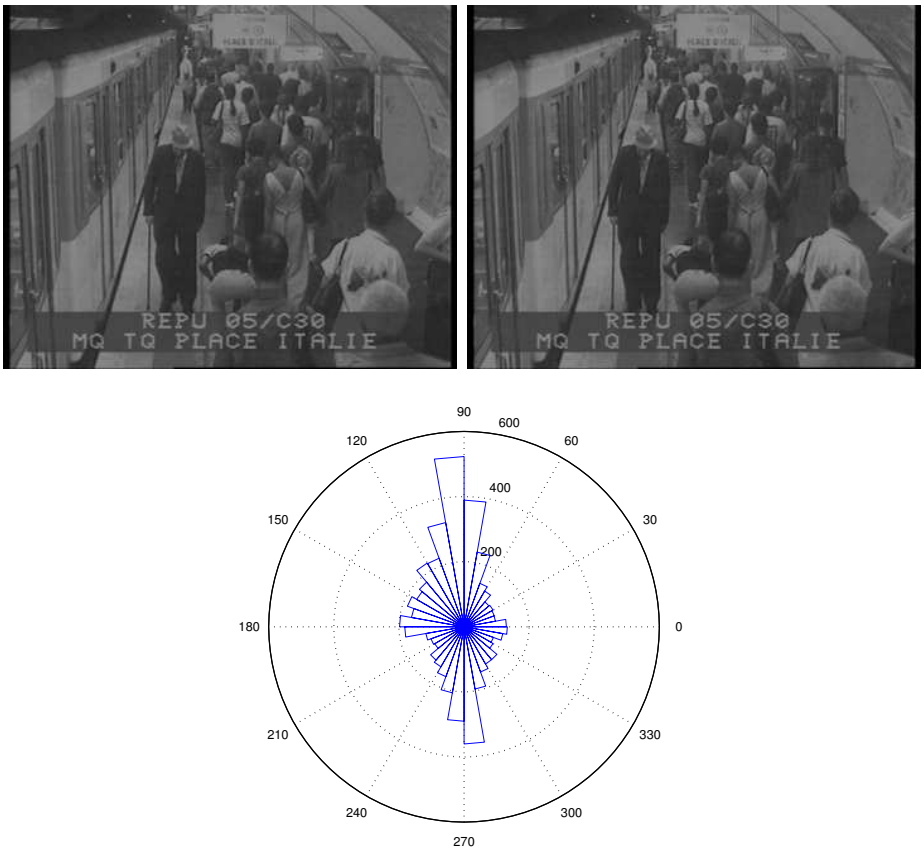


FIG. 5.4: Histogrammes des directions pour la méthode de Bernard (au dessus de l'histogramme, le couple d'images sur lesquels le flot optique est calculé).

5.5 Estimation de la densité d'une foule

Le but est maintenant d'estimer la densité d'une foule sur un quai de métro. L'information que nous nous proposons d'exploiter pour cela est la courbure des lignes de niveau d'une image. On rappelle que sous certaines conditions mathématiques sur l'image u (que $u : \mathbb{R}^2 \to \mathbb{R}$ soit une fonction semi-continue supérieurement), une ligne de niveau est définie comme la frontière d'une composante connexe d'un ensemble de niveau de u.

$$ll(\lambda)(x,y) = \partial(cc(\chi_\lambda(u))(x,y)) \text{ où } \chi_\lambda(u) = \{(x,y);\ u(x,y) \geq \lambda\}.$$

Soit une courbe C^1 $s \mapsto \mathbf{X}(s) = (x(s), y(s))$ paramétrée par son abscisse curviligne. On définit classiquement la courbure κ par un calcul différentiel dans les

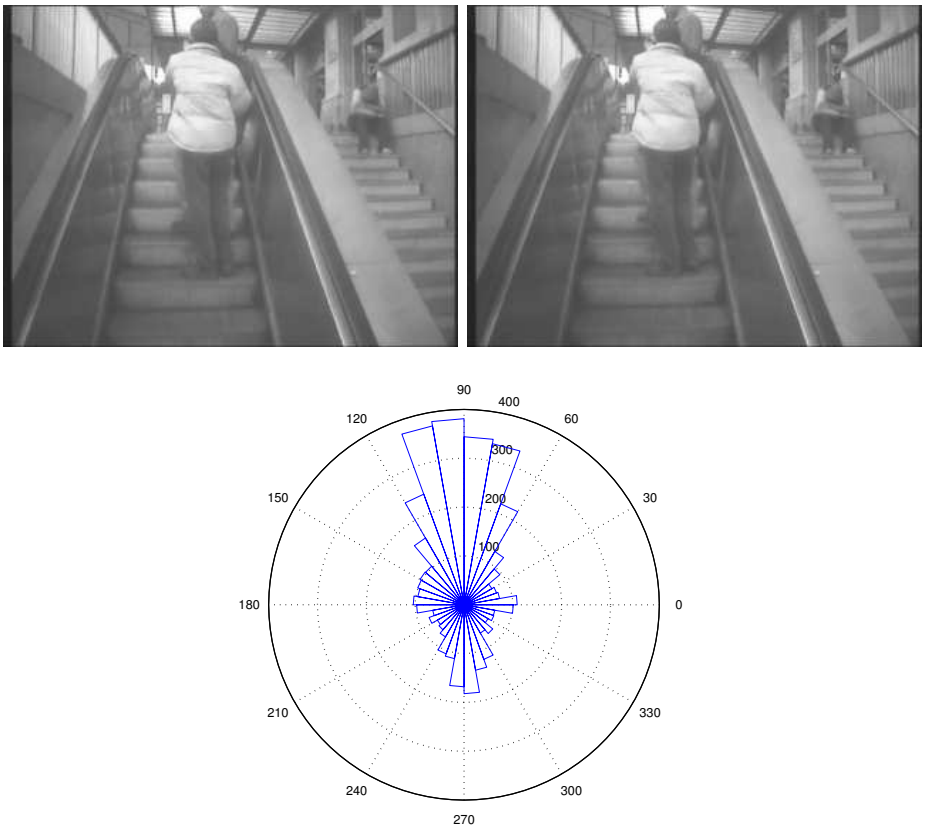


FIG. 5.5: Histogrammes des directions pour la méthode de Bernard entre la 520^{e} et la 521^{e} image de la séquence.

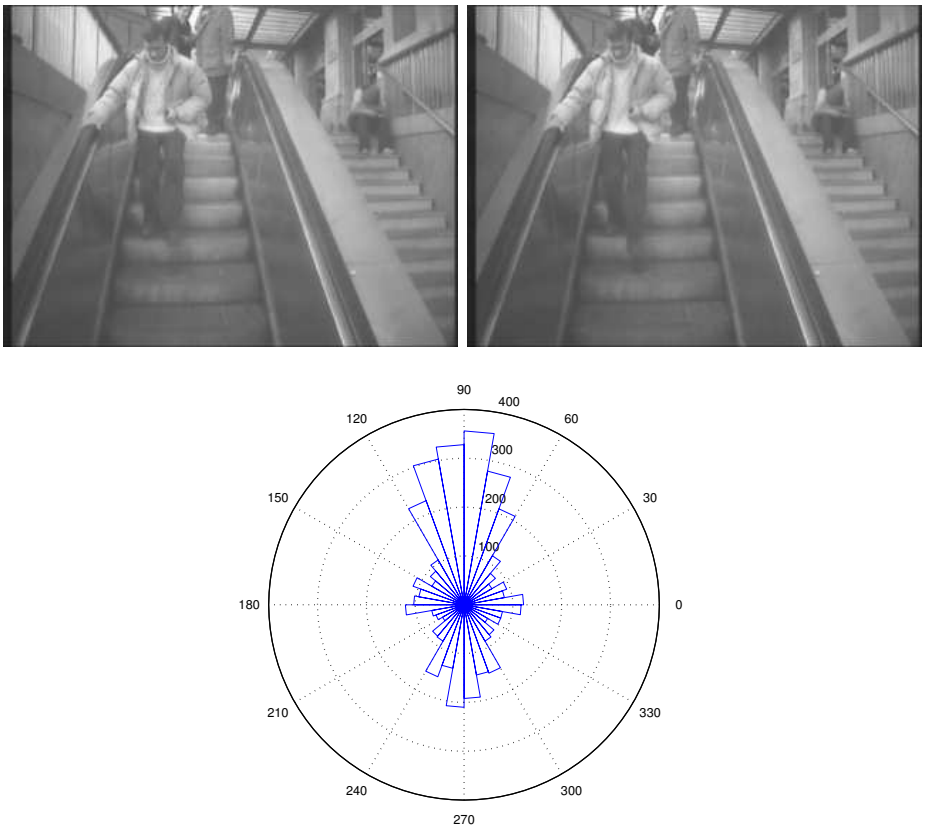


FIG. 5.6: Histogrammes des directions pour la méthode de Bernard entre la 580^{e} et la 581^{e} image de la séquence.

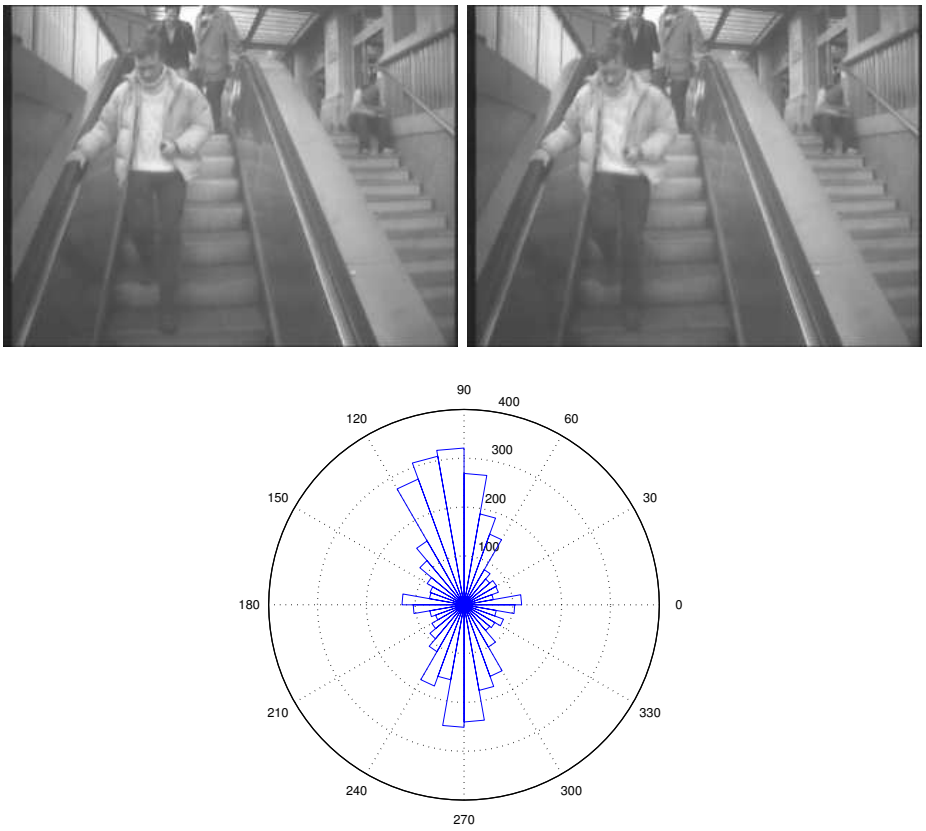


FIG. 5.7: Histogrammes des directions pour la méthode de Bernard entre la 599^{e} et la 600^{e} image de la séquence.

(a) 200^{e} image

(b) 480^{e} image

(c) 600^{e} image

(d) 900^{e} image

FIG. 5.8: Quatre moments clés de la séquence.

FIG. 5.9: En haut, deux images pour lesquelles une alarme correcte est générée ; en bas, deux images de fausse alarme.

coordonnées de Fréchet ($\mathbf{t} = \dfrac{\mathrm{d}\mathbf{X}}{\mathrm{d}s}, \mathbf{n} = \mathbf{t}^{\perp}$)

$$\frac{\mathrm{d}\mathbf{t}}{\mathrm{d}s} = -\kappa\mathbf{n}.$$

A partir de là, on peut aisément prouver

$$\kappa(s) = det(\mathbf{X}'(s), \mathbf{X}''(s)).$$

Dans le cas où la courbe paramétrée est une ligne de niveau de u, on a l'expression suivante, qui donne la courbure en fonction de u :

$$\kappa(s) = \mathrm{curv}(u)(x(s), y(s)) = \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) \tag{5.3}$$

Nous allons utiliser la courbure pour caractériser le remplissage d'un quai de métro : dans le cas d'une foule dense, la valeur absolue de la courbure dans l'image devrait être plus grande que dans le cas d'un quai plus ou moins vide, en raison de la courbure non négligeable induite par le profil d'une personne. Mais si nous tentons de calculer la courbure en tout point d'une image par (5.3) de manière brute, le résultat va être très irrégulier en raison de la grande irrégularité des lignes de niveau. Heureusement, on peut lisser les lignes de niveau de façon à les rendre plus régulières tout en gardant les principales caractéristiques de l'image. On applique alors l'une des deux équations d'évolution suivantes à l'image :

$$\frac{\partial u}{\partial t} = |Du|\mathrm{curv}(u)$$

(Mean Curvature Motion)

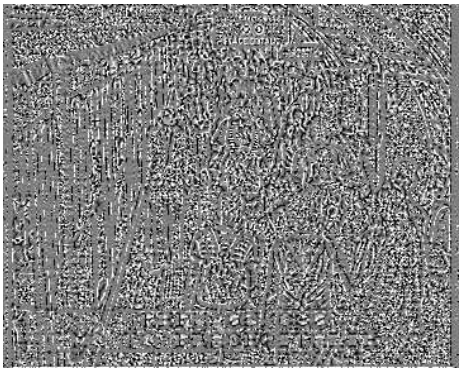
$$\frac{\partial u}{\partial t} = |Du|\mathrm{curv}(u)^{1/3}$$

(Affine Morphological Scale Space). Ces EDP ont été découvertes dans [AGLM] où les auteurs cherchaient des EDP vérifiant une certaine axiomatique.
Ces deux équations aux dérivées partielles régularisent les lignes de niveau principalement aux points de très forte courbure alors que les points de faible courbure évoluent plus lentement. L'algorithme pratique repose sur la discrétisation de l'EDP que l'on écrira sous la forme $\dfrac{\partial u}{\partial t} = |Du|\mathrm{curv}(u)^{\gamma}$: on a $u^{n+1} - u^n = \delta t(|Du^n|\mathrm{curv}(u^n)^{\gamma})$.
On obtient ainsi une suite d'images $(u^n)_{n=0}^{n_{end}}$ progressivement lissées ; et pour chacune de ces images, on calcule la valeur moyenne de la courbure en valeur absolue sur toute l'image

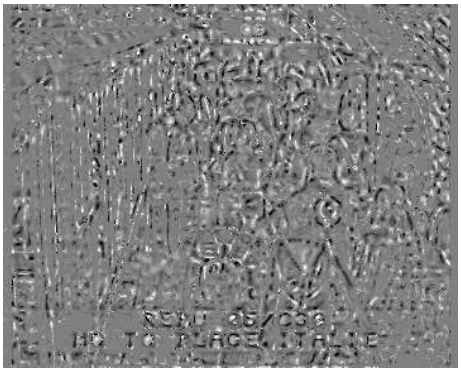
$$|\mathrm{curv}|_T(u^n) = \frac{1}{|\Omega|}\sum_{x,y\in\Omega} |\mathrm{curv}(u^n)(x,y)|$$

On aurait pu choisir d'autres critères plus généraux utilisant eux aussi l'information de courbure. Au lieu de considérer le premier moment de $|curv(u)|$, on pourrait considérer son histogramme et caractériser les périodes où la foule a une densité peu variable en fonction du résultat donné par la distance entre l'histogramme et celui de chacune des autres images. Néanmoins, le critère utilisant le premier moment nous a paru convenir dans nos expériences et permet d'obtenir des résultats quasiment en temps réel, nous en sommes donc resté à ce critère.
En pratique, afin de décider si on a à faire à une foule dense, moyenne ou faible, l'utilisateur se fixe deux références : il fait calculer $|\text{curv}|_T(u^n)$ à un moment où le quai de métro est quasiment vide et à un moment où la densité de la foule est quasiment à son maximum. On interpole ensuite linéairement (l'interpolation linéaire est justifiée par nos expériences, voir plus bas) ces deux valeurs et on se fixe des seuils correspondant au changement densité faible/densité moyenne, densité moyenne/densité forte. A partir de ce modèle linéaire, en fonction des valeurs de $|\text{curv}|_T(u^n)$ observées, on décide à laquelle des trois classes appartient l'image sur laquelle on travaille.
On présente en **figure 5.5** un résultat de ce calcul sur plusieurs images de la séquence du quai. Nous appliquons d'abord un masque pour s'intéresser seulement au quai, l'image ainsi modifiée est nulle en dehors du quai et inchangée sur le domaine du quai par rapport à l'image de départ. Si des zones des images à traiter sont sans intérêt pour le problème, on peut ainsi les rejeter et faire les calculs de courbure uniquement sur les zones d'intérêt. Ici, nous avons calculé la courbure totale en valeur absolue toutes les 10 images de la séquence, ce qui nous fait 80 images.

(a) Image sur laquelle on applique la MCM ou AMSS



(b) Courbure de u_1 ($t = t_1 = \delta t$, pour des raisons de visibilité, on applique le changement de contraste $k \mapsto 100k + 128$)



(c) Courbure de $u_{n_{end}}$ ($t = t_{end}$, même remarque)

FIG. 5.10: Deux résultats de courbure après lissage de l'image 5.10(a) par MCM ou AMSS (pas de temps : $\delta t = 0.1$, $t_{end} = 3.125$)

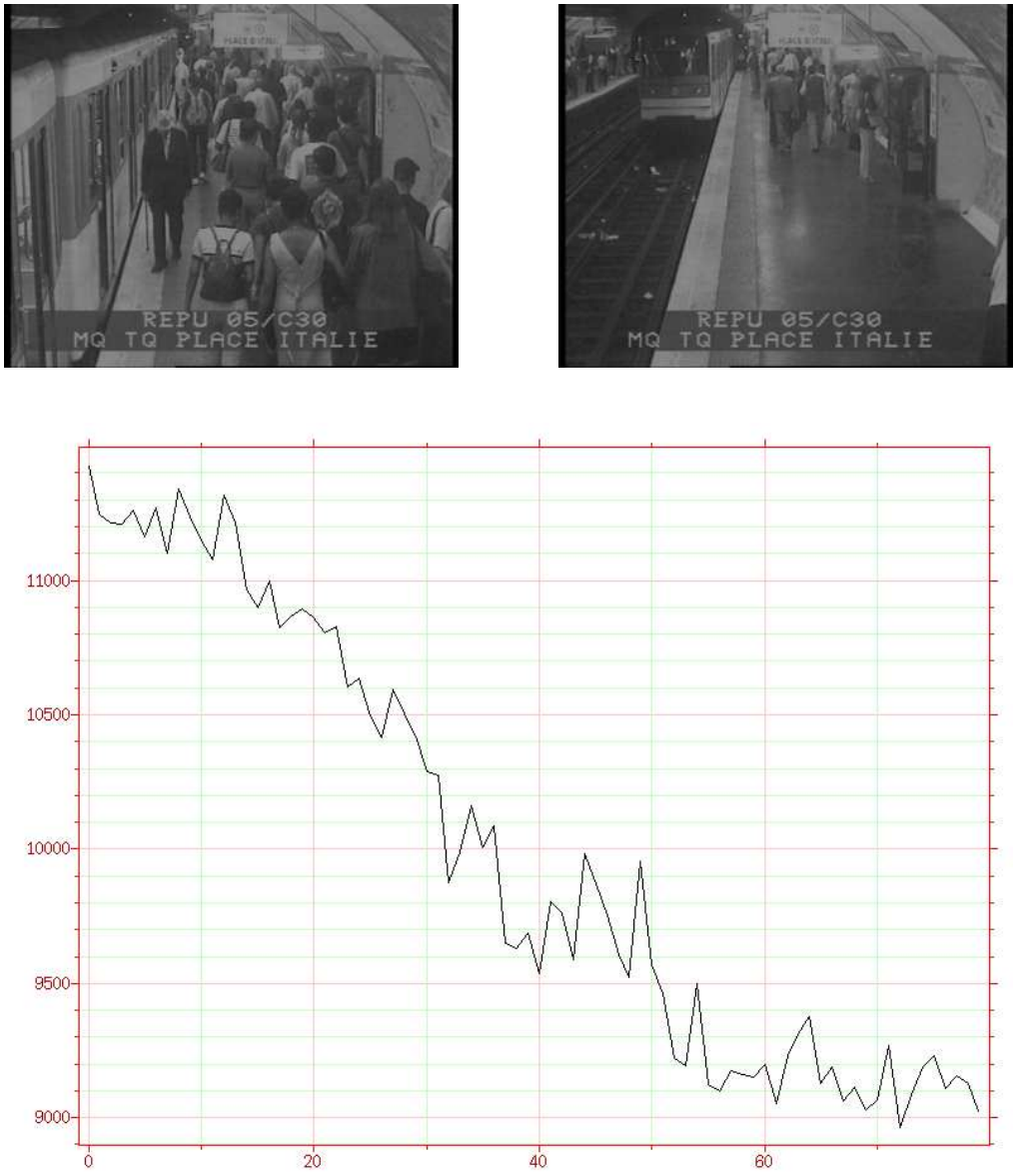


FIG. 5.11: Evolution de la courbure totale en valeur absolue (quantité évaluée toutes les 10 images). Les premières valeurs correspondent au quai très rempli (10^{e}image à gauche) et les dernières au quai presque vide (400^{e}image à droite).

Conclusion

Dans les trois premiers chapitres de cette thèse, nous avons présenté plusieurs méthodes de segmentation d'images fixes et d'objets en mouvement dans une vidéo.
Nous avons présenté une nouvelle méthode basée sur une détermination *a priori* par la méthode de Weickert et Schnörr. Cette méthode s'applique à des séquences filmées à caméra fixe (si la caméra est en mouvement, on recale donc les images) et où les objets ont un mouvement assez rigide (voitures, bâteaux), bien que nous ayons obtenu d'assez bons résultats sur une séquence avec des mouvements non rigides.
Les autres méthodes que nous avons présentées (méthode d'Aubert–Barlaud–Jehan-Besson) utilisent en général la donnée du fond. Elles sont donc tributaires de la qualité du fond comme nous sommes tributaires de la qualité du flot optique dans notre méthode. Toutefois lorsque le fond est de qualité correcte, les résultats de ces méthodes sont très bons et améliorent même le résultat que l'on obtiendrait simplement en seuillant la différence entre l'image et le fond.
Pour ce genre de méthodes, nous avons vu qu'il est intéressant, notamment pour des raisons de temps de calcul, d'utiliser l'algorithme ICM. Pour cela, nous avons mis en évidence les liens qui existent entre les modèles continus et les modèles discrets en segmentation d'images et de vidéo. Pour la discrétisation du terme de longueur, nous nous sommes servis de celle introduite par Chan et Song en la justifiant dans un cadre markovien. Nous avons appliqué l'ICM avec cette nouvelle discrétisation sur un modèle de suivi d'objets récemment introduit par Aron, Mansouri et Mitiche.
D'une manière générale, nous avons vu que nous pouvions appliquer l'ICM dans le cas de "descripteurs" ne dépendant pas de la région recherchée. Initialement, Chan et Song l'avaient appliqué à la fonctionnelle de Mumford-Shah/Chan-Vese où les descripteurs dépendent de la région. Nous pourrions donc nous demander jusqu'à quel point la méthode est généralisable au cas où les descripteurs dépendent de la région. Les travaux de Chan et Song permettent au moins d'envisager de pouvoir traiter le cas de descripteurs faisant intervenir des moments d'ordre au plus égal à deux.
Dans la dernière partie de cette thèse, nous avons proposé des solutions pour deux problèmes concrets de vidéo surveillance qui nous ont été soumis. Il s'agit

d'avoir une estimation de la densité d'une foule sur un quai de métro et de détecter des comportements "anormaux" de personnes. Pour le premier problème, nous avons proposé un algorithme très simple basé sur la courbure algébrique des images. Cet algorithme tourne presque en temps réel, bien que cette contrainte ne nous ait pas été vraiment posée.
Pour le deuxième problème, nous avons utilisé des modèles simples de flot optique, car la contrainte du temps réel est ici très importante. Nous nous en approchons mais nous n'y sommes pas encore. Les résultats sont toutefois encourageants, puisqu'ils sont statistiquement en accord avec l'observation subjective de la scène par un tiers. Il est à noter que l'algorithme fonctionne à partir d'un seuil choisi *ad hoc*, mais une fois cette valeur trouvée, elle peut servir de référence pour d'autres séquences où la même scène est filmée.

Annexe A

Théorie de la dérivation de domaine

Dans cette partie, nous présentons plusieurs résultats de la théorie dite des *dérivées de forme* ou *dérivées de domaine* introduite par Sokolowski et Zolésio dans [SZ] et qui a donnée lieu à la monographie de Delfour et Zolésio [DZ]. Ces résultats sont appliqués aux fonctionnelles rencontrées dans les premiers chapitres de cette thèse. Il s'agit d'en calculer les dérivées et pour cela, comme le nom de la théorie l'indique, on définit des nouvelles notions de dérivée généralisant les dérivées classiques.
Au niveau des notations, nous suivrons plutôt Ring et Hintermüller ([HR]) qui reprennent les différents calculs de [DZ]. La dérivée particulaire de ϕ sera donc notée $\dot{\phi}$ et la dérivée de forme ϕ'. Le produit scalaire dans $\mathbb{R}^n$ sera noté indifféremment $a \cdot b$ ou $< a, b >$.

Définition 2 *Soit $V : \mathbb{R}^2 \to \mathbb{R}^2$ un champ de vecteurs de classe C^k à support compact et $\mathbf{x} \in \mathbb{R}^2$. On définit le flot, $T_t : \mathbb{R}^2 \to \mathbb{R}^2, \mathbf{x} \to \mathbf{X}(t)$[1] où $\mathbf{X}(t)$ désigne l'unique solution de*

$$\left\{ \begin{array}{rcl} \mathbf{X}'(t) & = & V(\mathbf{X}(t)) \\ \mathbf{X}(0) & = & \mathbf{x} \end{array} \right. . \tag{A.1}$$

Pour un contour Γ et un ouvert Ω, on définit leurs image par T_t

$$\begin{array}{rcl} \Gamma_t & = & \{T_t(\mathbf{x}),\, \mathbf{x} \in \Gamma\} = T_t(\Gamma) \\ \Omega_t & = & \{T_t(\mathbf{x}),\, \mathbf{x} \in \Omega\} = T_t(\Omega). \end{array}$$

Comme $V \in C_0^k(\mathbb{R}^2, \mathbb{R}^2)$, T_t est de classe C^k et donc Γ_t et Ω_t ont les mêmes propriétés de régularité que Γ et Ω.

[1] Bien que le flot dépende de V, on notera plutôt T_t que $T_t(V)$ pour ne pas alourdir les notations.

Définition 3 *Soit $J : \mathcal{G} \to \mathbb{R}$ où $\mathcal{G}$ est un ensemble de contours (resp. un ensemble de domaines du plan).*
On définit la dérivée Eulérienne *de J en Γ (resp. Ω) dans la direction donnée par le champ V par*

$$\begin{array}{ll} & dJ(\Gamma;V) = \lim_{t\downarrow 0} \frac{1}{t}\left(J(\Gamma_t) - J(\Gamma)\right) \\ (resp. & dJ(\Omega;V) = \lim_{t\downarrow 0} \frac{1}{t}\left(J(\Omega_t) - J(\Omega)\right)\big) \end{array} \tag{A.2}$$

Définition 4 *Soit B un espace de Banach de champs de vecteurs de perturbation.*
La fonctionnelle J est dite différentiable au sens de la forme *en Γ (resp. Ω) dans B si $dJ(\Gamma;V)$ (resp. $dJ(\Omega;V)$) existe pour tout $V \in B$ et si l'application $V \mapsto dJ(\Gamma;V)$ (resp. $V \mapsto dJ(\Omega;V)$) est linéaire et continue dans B.*

Soient Γ (*resp.* Ω) appartenant à un certain ensemble de formes $\mathcal{G}$. On définit $G(\Gamma)$ (resp. $G(\Omega)$) comme l'ensemble des fonctions définies sur Γ (resp. Ω). $G(\Gamma)$ (*resp.* $G(\Omega)$) est un espace de Banach.
Soient deux fonctions $\phi(\Omega) : \Omega \to \mathbb{R},\ \mathbf{x} \mapsto \phi(\Omega)(\mathbf{x})$ et $\psi(\Gamma) : \Gamma \to \mathbb{R},\ \mathbf{x} \mapsto \psi(\Gamma)(\mathbf{x})$, on définit leurs *dérivées particulaires* de la manière suivante.

Définition 5 *On reprend les notations de la Définition 2 avec $k = 1$. On pose $\psi^t = \psi(\Gamma_t) \circ T_t$ (resp. $\phi^t = \phi(\Omega_t) \circ T_t$), $\psi^0 = \psi(\Gamma)$ (resp. $\phi^0 = \phi(\Omega)$) et on suppose que $\psi^t \in G(\Gamma_t)$ (resp. $\phi^t \in G(\Omega_t)$) pour tout $t \in]0,T[$ pour un certain $T > 0$. Si la limite*

$$\dot{\psi}(\Gamma;V) = \lim_{t\downarrow 0} \frac{1}{t}\left(\psi^t - \psi^0\right) \quad \left(resp.\ \dot{\phi}(\Omega;V) = \lim_{t\downarrow 0} \frac{1}{t}\left(\phi^t - \phi^0\right)\right) \tag{A.3}$$

existe pour la topologie faible de $G(\Gamma)$ (resp. $G(\Omega)$), i.e. *(on note $G(\Gamma)'$ [resp. $G(\Omega)'$] le dual de $G(\Gamma)$ [resp. $G(\Omega)$])*

$$\begin{array}{ll} & \forall g \in G(\Gamma)', < \dot{\psi}(\Gamma;V), g >_{G(\Gamma)\times G(\Gamma)'} = \lim_{t\downarrow 0} < \frac{\psi^t - \psi^0}{t}, g >_{G(\Gamma)\times G(\Gamma)'} \\ (\text{resp.} & \forall g \in G(\Omega)', < \dot{\phi}(\Omega;V), g >_{G(\Omega)\times G(\Omega)'} = \lim_{t\downarrow 0} < \frac{\phi^t - \phi^0}{t}, g >_{G(\Omega)\times G(\Omega)'}) \end{array} \tag{A.4}$$

alors $\dot{\psi}(\Gamma;V)$ (resp. $\dot{\phi}(\Omega;V)$) est appelée la dérivée particulaire ou dérivée totale faible de ψ (resp. ϕ) dans la direction V.

On a le même énoncé pour la topologie forte.
L'appellation de dérivée particulaire est reprise de la mécanique des fluides où pour une fonction $f = f(\mathbf{x}, t)$ et $g(t) = f(\mathbf{x}(t), t)$, elle désigne la dérivée de g

$$\frac{Df}{Dt} = g'(t) = \frac{\partial f}{\partial t} + < \nabla_{\mathbf{x}} f, V >$$

où $V = \mathbf{x}'(t)$. De la même manière, si la fonction ϕ dépend de la forme et du point, la dérivée particulaire au sens que nous lui avons donné fait intervenir le flot de la forme et le flot du point par le champ de déformation V, ainsi il

apparaît que si la fonction ϕ ne dépend pas de Ω, alors sa dérivée particulaire est donnée par

$$\dot{\phi}(\Omega; V) = \dot{\phi}(V) =< \nabla\phi, V > .$$

Afin de respecter l'idée intuitive selon laquelle la dérivée par rapport à Ω d'une fonction indépendante de Ω vaut 0, on introduit la dérivée de forme qui est en quelque sorte l'analogue du terme $\dfrac{\partial f}{\partial t}$ de la dérivée particulaire en mécanique des fluides.

Définition 6 *On suppose que la dérivée particulaire $\dot{\phi}(\Omega; V)$ et l'expression $< \nabla\phi, V >$ sont bien définies dans $G(\Omega)$. On définit la dérivée de forme de ϕ en Ω dans la direction V par*

$$\phi'(\Omega; V) = \dot{\phi}(\Omega; V) - < \nabla\phi, V > \tag{A.5}$$

Pour définir la dérivée particulaire de ψ, on a besoin d'introduire son *gradient tangentiel* $\nabla_\Gamma\psi$. On se donne une extension aussi régulière que l'on veut $\tilde{\psi}$ de ψ à tout le domaine D, le gradient tangentiel est alors définit comme

$$\nabla_\Gamma\psi \triangleq \nabla\tilde{\psi}|_\Gamma - \frac{\partial\tilde{\psi}}{\partial n}\mathbf{n}.$$

Définition 7 *On suppose que la dérivée particulaire $\dot{\psi}(\Gamma; V)$ et l'expression $< \nabla_\Gamma\phi, V >$ sont bien définies dans $G(\Gamma)$. On définit la dérivée de forme de ψ en Γ dans la direction V par*

$$\psi'(\Gamma; V) = \dot{\psi}(\Gamma; V) - < \nabla_\Gamma\psi, V > |_\Gamma \tag{A.6}$$

Pour bien comprendre comment les notions de dérivées de forme et de dérivée particulaire sont nécessaires à ce qui suit, nous allons écrire la fonctionnelle

$$J(\Omega_t(V)) = \int_{\Omega_t} \phi \, d\mathbf{x} = \int_\Omega \phi \circ T_t \, J_t \, d\mathbf{x},$$

par la formule du changement de variable ($J_t = \det(DT_t)$). Pour calculer la dérivée Eulérienne de cette fonctionnelle, nous avons besoin de calculer la limite de $\frac{1}{t}\left(J(\Omega_t) - J(\Omega)\right)$. Or

$$\frac{1}{t}\left(J(\Omega_t) - J(\Omega)\right) = \int_\Omega \underbrace{\frac{1}{t}\left(\phi^t - \phi^0\right)}_{\to\dot{\phi}(\Omega;V)} \underbrace{J_t}_{\to J_0=\det(Id)=1} d\mathbf{x} + \int_\Omega \underbrace{\phi^0 \frac{1}{t}\left(J_t - J_0\right)}_{\to\phi\frac{\mathrm{d}J_t}{\mathrm{d}t}|_{t=0}} d\mathbf{x}.$$

La dérivée particulaire apparaît donc naturellement. D'après [DZ] (Th 4.1, chap. 8), le flot de J_t est donné par

$$\frac{\mathrm{d}J_t}{\mathrm{d}t} = \mathrm{div}(V) \circ T_t.$$

On utilise alors la relation $\phi'(\Omega;V) = \dot{\phi}(\Omega;V) - <\nabla\phi, V>$ pour faire apparaître $\phi'(\Omega;V)$ et regrouper les termes $\int_\Omega <\nabla\phi, V>\, d\mathbf{x}$ et $\int_\Omega \phi \,\mathrm{div}(V)\, d\mathbf{x}$ en la seule intégrale

$$\int_\Omega \mathrm{div}(\phi V)\, d\mathbf{x} = \int_\Gamma \phi <V, \mathbf{n}>\, d\mathcal{H}^1$$

par la formule de Stokes si Γ est Lipschitz. Il vient alors le théorème suivant.

Théorème 8 *Soit $\phi = \phi(\Omega) \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^2)$ telle que sa dérivée particulaire au sens faible L^1 $\dot{\phi}(\Omega;V)$ et sa dérivée de forme $\phi'(\Omega;V)$ existent. On suppose également que Ω est ouvert borné et $\Gamma = \partial\Omega$ est de classe C^1. Alors la fonctionnelle*

$$J(\Omega) = \int_\Omega \phi(\Omega, \mathbf{x})\, d\mathbf{x} \tag{A.7}$$

admet une dérivée eulérienne dans la direction V donnée par

$$dJ(\Omega;V) = \int_\Omega \phi'(\Omega;V)\, d\mathbf{x} + \int_\Gamma \phi <V, \mathbf{n}>\, d\mathcal{H}^1. \tag{A.8}$$

En faisant les mêmes hypothèses sur l'existence de la dérivée particulaire et la dérivée de forme mais cette fois avec $\phi \in W^{2,1}_{\mathrm{loc}}(\mathbb{R}^2)$ et Γ de classe C^2, la fonctionnelle définie par

$$J(\Gamma) = \int_\Gamma \phi(\Omega, \mathbf{x})|_\Gamma\, d\mathcal{H}^1, \tag{A.9}$$

admet comme dérivée eulérienne dans la direction V

$$dJ(\Gamma;V) = \int_\Gamma \phi'(\Omega;V)|_\Gamma\, d\mathcal{H}^1 + \int_\Gamma (\frac{\partial\phi}{\partial n} + \kappa\phi) <V, \mathbf{n}>\, d\mathbf{x} \tag{A.10}$$

où κ désigne la courbure de Γ et $\dfrac{\partial\phi}{\partial n} =< \nabla\phi, \mathbf{n} >$.

Remarque : il existe un résultat sur les fonctionnelles $J(\Gamma) = \int_\Gamma \psi(\Gamma, \mathbf{x})\, d\mathcal{H}^1$, mais dans la plupart des cas, les fonctionnelles utilisées en traitement d'images font intervenir en intégrande des fonctions qui sont définies sur tout le domaine Ω et non spécifiquement sur sa frontière. Pour cette raison, nous nous en sommes tenu à cette présentation des résultats de dérivées eulériennes.

A.1 Dérivées d'énergies classiques

A.1.1 Longueur, aire et géodésiques

Le Théorème 8 permet d'exprimer les dérivées de forme de nombreuses fonctionnelles utilisées en segmentation d'images par contours actifs, ainsi la fonctionnelle *longueur* de Γ

$$\mathcal{L}(\Gamma) = \int_\Gamma d\mathcal{H}^1$$

admet une dérivée dans la direction V donnée simplement par

$$\boxed{d\mathcal{L}(\Gamma;V) = \int_\Gamma \kappa < V, \mathbf{n} > \, d\mathcal{H}^1.}$$

La direction de plus grande descente est donc donnée par

$$\frac{\partial \Gamma}{\partial t} = V = -\kappa \mathbf{n}$$

(On rappelle que d'après les conventions de [DZ], $\mathbf{n}$ désigne la normale extérieure à Γ, cette EDP induit bien une évolution vers l'*intérieur* de Γ à cause du signe moins).
La fonctionnelle *aire du domaine intérieur* à $\Gamma = \partial\Omega$

$$\mathcal{A}(\Omega) = \int_\Omega d\mathbf{x}$$

admet comme dérivée

$$\boxed{d\mathcal{A}(\Omega;V) = \int_\Gamma < V, \mathbf{n} > \, d\mathcal{H}^1.}$$

La direction de plus grande descente est cette fois

$$\frac{\partial \Gamma}{\partial t} = V = -\mathbf{n}.$$

De même, pour l'énergie des contours actifs géodésiques de Caselles, Kimmel et Sapiro

$$\mathcal{E}_{CKS}(\Gamma) = \int_\Gamma g \, d\mathcal{H}^1$$

où $g = g(\mathbf{x})$, la dérivée de forme est donnée par

$$\boxed{d\mathcal{E}_{CKS}(\Gamma;V) = \int_\Gamma (\kappa g + \nabla g \cdot \mathbf{n}) < V, \mathbf{n} > \, d\mathcal{H}^1.}$$

qui est bien le résultat usuel.

A.1.2 Energies impliquant moyennes et variances

On reprend les travaux d'Aubert, Barlaud et Jehan-Besson que nous avons décrits en détail dans un autre chapitre. Dans [ABJB2], ces auteurs cherchent à segmenter des objets présentant une certaine homogénéité de la variance. Ils considèrent donc la fonctionnelle

$$\mathcal{J}_{ABJB}(\Omega) = \int_\Omega \phi(\sigma^2_\Omega)\, \mathbf{x} + \int_{\Omega^c} \phi(\sigma^2_{\Omega^c})\, \mathbf{x} + \lambda \int_\Gamma d\mathcal{H}^1 \qquad \text{(A.11)}$$

où

$$\mu_\Omega = \frac{\int_\Omega I(\mathbf{x})\, d\mathbf{x}}{|\Omega|}$$

et

$$\sigma^2_\Omega = \frac{\int_\Omega (I(\mathbf{x}) - \mu_\Omega)^2 \, d\mathbf{x}}{|\Omega|}.$$

D'après le Théorème 8, pour calculer la dérivée dans la direction V de $\mathcal{J}_{ABJB}$, on a besoin de calculer la dérivée de forme de $\phi(\sigma^2_\Omega)$. Comme seul σ^2_Ω ne dépend que de Ω, il apparaît que

$$\left(\phi(\sigma^2_\Omega)\right)'(\Omega; V) = \phi'(\sigma^2_\Omega)\left(\sigma^2_\Omega\right)'(\Omega; V).$$

Il nous reste donc à calculer $\left(\sigma^2_\Omega\right)'(\Omega; V)$. Comme σ^2_Ω est quotient de deux fonctionnelles dépendant de Ω, sa dérivée se calcule comme une dérivée de quotient usuelle (la démonstration est identique). Par ailleurs, les dérivées de forme du numérateur et du dénominateur sont exactement les dérivées de ces quantités en tant que fonctionnelles dans la direction V. On a donc

$$\left(\sigma^2_\Omega\right)'(\Omega; V) = \frac{\int_\Omega \left((I(\mathbf{x}) - \mu_\Omega)^2\right)'(\Omega; V)\, d\mathbf{x} + \int_\Gamma (I(\mathbf{x}) - \mu_\Omega)^2 < V, \mathbf{n} > \, d\mathcal{H}^1}{|\Omega|} - \sigma^2_\Omega \frac{\int_\Gamma < V, \mathbf{n} > \, d\mathcal{H}^1}{|\Omega|^2}. \quad \text{(A.12)}$$

De manière évidente, on a $\left((I(\mathbf{x}) - \mu_\Omega)^2\right)'(\Omega; V) = -2(I(\mathbf{x}) - \mu_\Omega)\mu'_\Omega(\Omega; V)$. Il ne nous reste plus qu'à calculer $\mu'_\Omega(\Omega; V)$. Exactement par le même raisonnement, on obtient

$$\mu'_\Omega(\Omega; V) = \frac{\int_\Gamma (I - \mu_\Omega) < V, \mathbf{n} > \, d\mathcal{H}^1}{|\Omega|}.$$

Finalement, $\mu'_\Omega(\Omega; V)$ ne dépend pas du point, il peut donc être sorti de la première intégrale dans A.12, ce qui rend ce terme nul. En définitive,

$$\left(\sigma^2_\Omega\right)'(\Omega; V) = \frac{\int_\Gamma [(I - \mu_\Omega)^2 - \sigma^2_\Omega] < V, \mathbf{n} > \, d\mathcal{H}^1}{|\Omega|}. \quad \text{(A.13)}$$

On peut donc en déduire la dérivée de $\mathcal{J}_{ABJB}$ dans la direction V. D'après le théorème 1, on a

$$\begin{aligned} d\mathcal{J}_{ABJB}(\Omega; V) &= \int_\Omega (\phi(\sigma^2_\Omega))'(\Omega; V)\, d\mathbf{x} + \int_{\Omega^c} (\phi(\sigma^2_{\Omega^c}))'(\Omega; V)\, d\mathbf{x} \\ &\quad + \int_\Gamma \left(\phi(\sigma^2_\Omega) - \phi(\sigma^2_{\Omega^c}) + \lambda\kappa\right) < V, \mathbf{n} > \, d\mathcal{H}^1 \end{aligned}$$

$$\boxed{\begin{aligned} d\mathcal{J}_{ABJB}(\Omega; V) &= \int_\Gamma \left(\phi(\sigma^2_\Omega) + \phi'(\sigma^2_\Omega)[(I - \mu_\Omega)^2 - \sigma^2_\Omega]\right. \\ &\quad \left. - \phi(\sigma^2_{\Omega^c}) - \phi'(\sigma^2_{\Omega^c})[(I - \mu_{\Omega^c})^2 - \sigma^2_{\Omega^c}] + \lambda\kappa\right) < V, \mathbf{n} > \, d\mathcal{H}^1 \end{aligned}}$$

A.2 Dérivée de $\boldsymbol{J(\Gamma) = \int_\Gamma g(\nabla u \cdot \mathbf{n})\, d\mathcal{H}^1}$

Considérons la fonctionnelle

$$J(\Gamma) = \int_\Gamma g(\nabla u \cdot \mathbf{n})\, d\mathcal{H}^1. \tag{A.14}$$

Cette fonctionnelle représente une mesure de l'alignement entre la normale à la courbe et le vecteur gradient d'une fonction u. Il ne pose pas de problème de considérer l'alignement avec un champ de vecteur ne dérivant pas *a priori* d'une fonction comme c'est le cas de ∇u, on peut donc remplacer ∇u par n'importe quel champ de vecteurs A (et donc Δu par $\mathrm{div}(A)$) dans toute la suite.
La dérivée dans la direction V de la fonctionnelle nous donne

$$dJ(\Gamma; V) = \int_\Gamma \left(g(\nabla u \cdot \mathbf{n})\right)'(\Gamma; V)|_\Gamma \, d\mathcal{H}^1 + \int_\Gamma \left(\frac{\partial}{\partial n}\left(g(\nabla u \cdot \mathbf{n})\right) + g(\nabla u \cdot \mathbf{n})\kappa\right) < V, \mathbf{n} > \, d\mathcal{H}^1.$$

Calcul de $\boldsymbol{\left(g(\nabla u \cdot \mathbf{n})\right)'(\Gamma; V)}$
Comme il n'y a que $\mathbf{n}$ qui dépend de la forme, on a

$$\left(g(\nabla u \cdot \mathbf{n})\right)'(\Gamma; V) = g'(\nabla u \cdot \mathbf{n}) \nabla u \cdot \mathbf{n}'$$

Comme Ring et Hintermüller l'ont souligné dans [HR], pour un tel calcul, il est intéressant d'utiliser la fonction distance signée à Γ que l'on notera b_Γ. Si on note

$$d_\Omega(\mathbf{x}) \triangleq \inf_{y \in \Omega} \|\mathbf{x} - \mathbf{y}\|,$$

alors

$$b_\Gamma(\mathbf{x}) = d_\Omega(\mathbf{x}) - d_{\mathbb{R}^2 \setminus \Omega}(\mathbf{x}).$$

En effet, on considère alors que $\mathbf{n} = \nabla b_\Gamma|_\Gamma$ (on se référera à [HR] pour toutes les explications techniques).
On calcule $\mathbf{n}' = (\nabla b_\Gamma)'$, mais d'après [HR] (3.13),

$$(\nabla b_\Gamma)' = \nabla_\Gamma b_\Gamma' = -\nabla_\Gamma \left(< V, \mathbf{n} >\right).$$

Finalement, on a

$$\left(g(\nabla u \cdot \mathbf{n})\right)'(\Gamma; V) = -g'(\nabla u \cdot \mathbf{n}) \nabla u \cdot \nabla_\Gamma \left(< V, \mathbf{n} >\right) \tag{A.15}$$

Calcul de $\boldsymbol{\dfrac{\partial}{\partial n}\left(g(\nabla u \cdot \mathbf{n})\right)}$
On utilise encore la fonction distance. On a

$$\frac{\partial}{\partial n}\left(g(\nabla u \cdot \mathbf{n})\right) = \nabla\left(g(\nabla u \cdot \mathbf{n})\right) \cdot \mathbf{n} = \nabla\left(g(\nabla u \cdot \nabla b_\Gamma)\right) \cdot \nabla b_\Gamma|_\Gamma$$

Or

$$\nabla\left(g(\nabla u\cdot\nabla b_\Gamma)\right)=g'(\nabla u\cdot\nabla b_\Gamma)(D^2u\,\nabla b_\Gamma+\nabla u D^2 b_\Gamma).$$

On en déduit

$$\frac{\partial}{\partial n}\left(g(\nabla u\cdot\mathbf{n})\right)=g'(\nabla u\cdot\nabla b_\Gamma)(D^2u\,\nabla b_\Gamma\cdot\nabla b_\Gamma+D^2b_\Gamma\nabla b_\Gamma\cdot\nabla b_\Gamma)|_\Gamma.$$

Mais le module du gradient de la fonction distance est égal à 1, et en prenant le gradient de $|\nabla b_\Gamma|^2=1$, on peut en déduire comme dans [HR] que $D^2b_\Gamma\nabla b_\Gamma=0$. Finalement, on a donc

$$\frac{\partial}{\partial n}\left(g(\nabla u\cdot\mathbf{n})\right)=g'(\nabla u\cdot\mathbf{n})D^2u\,\mathbf{n}\cdot\mathbf{n} \tag{A.16}$$

Des calculs précédents, on tire l'expression de la dérivée dans la direction V

$$\begin{aligned}dJ(\Gamma;V)=-\int_\Gamma g'(\nabla u\cdot\mathbf{n})\nabla u\cdot\nabla_\Gamma\left(<V,\mathbf{n}>\right)\,d\mathcal{H}^1\\ +\int_\Gamma(g'(\nabla u\cdot\mathbf{n})D^2u\,\mathbf{n}\cdot\mathbf{n}+g(\nabla u\cdot\mathbf{n})\kappa)<V,\mathbf{n}>\,d\mathcal{H}^1.\end{aligned}$$

Afin de simplifier le premier terme, on peut utiliser la *formule de Green tangentielle* (*cf.* [DZ] (5.27, chap. 8))

$$\int_\Gamma f\mathrm{div}_\Gamma g+\nabla_\Gamma f\cdot g\,d\mathcal{H}^1=\int_\Gamma\kappa fg\cdot\mathbf{n}\,d\mathcal{H}^1. \tag{A.17}$$

On obtient donc ($\frac{\partial u}{\partial n}=\nabla u\cdot\mathbf{n}$)

$$\boxed{\begin{array}{rcl}dJ(\Gamma;V) &=& \int_\Gamma\left\{\mathrm{div}_\Gamma\left[g'\left(\frac{\partial u}{\partial n}\right)\nabla u\right]-g'\left(\frac{\partial u}{\partial n}\right)\frac{\partial u}{\partial n}\kappa+g\left(\frac{\partial u}{\partial n}\right)\kappa\right.\\ &&\left.+g'\left(\frac{\partial u}{\partial n}\right)D^2u\,\mathbf{n}\cdot\mathbf{n}\right\}<V,\mathbf{n}>\,d\mathcal{H}^1\\ &=& \int_\Gamma\left\{\nabla_\Gamma\left(g'\left(\frac{\partial u}{\partial n}\right)\right)\cdot\nabla u+g'\left(\frac{\partial u}{\partial n}\right)\Delta u\right.\\ &&\left.+h\left(\frac{\partial u}{\partial n}\right)\kappa\right\}<V,\mathbf{n}>\,d\mathcal{H}^1,\end{array}}$$

avec $h(t)=g(t)-tg'(t)$. En effet,

$$\mathrm{div}_\Gamma\left[g'\left(\frac{\partial u}{\partial n}\right)\nabla u\right]=\nabla_\Gamma\left(g'\left(\frac{\partial u}{\partial n}\right)\right)\cdot\nabla u+g'\left(\frac{\partial u}{\partial n}\right)\mathrm{div}_\Gamma(\nabla u),$$

et $\mathrm{div}_\Gamma(\nabla u)=\mathrm{div}(\nabla u)-D^2u\,\mathbf{n}\cdot\mathbf{n}=\Delta u-D^2u\,\mathbf{n}\cdot\mathbf{n}$ d'après (5.19, chap. 8) dans [DZ].
A noter que pour $g(t)=t^2$, en utilisant la relation $\mathrm{div}_\Gamma f-\kappa f\cdot\mathbf{n}=\mathrm{div}_\Gamma f_\Gamma$, avec

$f_\Gamma = f - f \cdot \mathbf{n}\ \mathbf{n}$, on obtient bien la formule (5.29, chap. 8) donnée dans [DZ]. A partir de là, on peut en déduire la dérivée de l'énergie de Kimmel-Bruckstein

$$\mathcal{E}_{KB}(\Gamma) = \frac{J(\Gamma)}{\mathcal{L}(\Gamma)}.$$

On peut alors appliquer les règles usuelles sur le quotient de deux fonctionnelles pour trouver l'expression générale de la dérivée

$$d\mathcal{E}_{KB}(\Gamma; V) = \frac{dJ(\Gamma; V)}{\mathcal{L}(\Gamma)} - \frac{J(\Gamma; V) d\mathcal{L}(\Gamma; V)}{\mathcal{L}(\Gamma)^2}.$$

Annexe B

Documentation des modules MegaWave utilisés

Le but de ce chapitre est de documenter les implémentations que nous avons faites des différents algorithmes que nous avons décris dans les chapitres précédents, afin d'en faciliter l'utilisation. Nous avons utilisé l'environnement de développement MegaWave créé par J. Froment en 1993 au CEREMADE de l'Université Paris IX-Dauphine. MegaWave est à présent maintenu et mis à jour par J. Froment et L. Moisan. Comme la mise à jour est progressive, le lecteur est invité à nous contacter s'il veut disposer des implémentations qui n'ont pas été intégrées dans la version courante de MegaWave.

○Name

icm Iterated Conditional Modes algorithm

○Command Synopsis

icm [-p *prec*] *flag sweep alpha mu im input_mask output_mask*

-p *prec* : precision between means for Mumford-Shah energy
flag : flag=0 : threshold energy, flag=1 : Mumford-Shah/ Chan-Vese, flag=2 : Tsai-Yezzi-Willsky
sweep : Number of sweeps
alpha : threshold parameter in threshold energy
mu : regularization parameter
im : image the algorithm is to be applied
input_mask : -1/+1 input mask image
output_mask : -1/+1 output mask image

○Function Summary

void icm (flag , prec , sweep , alpha , mu , im , input_mask , output_mask)
int flag ;
float *prec ;
int sweep ;
float alpha , mu ;
Fimage im ;
Fimage input_mask , output_mask ;

○Description

This module implements the Iterated Conditional Modes algorithm (ICM) for the 3 energies

$$E(\phi) = \sum_{i,j} (f(i,j) - c_1(\phi))^2 \, H(\phi(i,j)) + (f(i,j) - c_2(\phi))^2 \, (1 - H(\phi(i,j))) + \mu \sum_{i,j} \underbrace{\sqrt{(H(\phi(i+1,j)) - H(\phi(i,j)))^2 + (H(\phi(i,j+1)) - H(\phi(i,j)))^2}}_{\triangleq l(i,j)}. \tag{B.1}$$

$$E(\phi) = -\frac{1}{2}\left(c_1(\phi) - c_2(\phi)\right)^2 + \mu \sum_{i,j} \underbrace{\sqrt{(H(\phi(i+1,j)) - H(\phi(i,j)))^2 + (H(\phi(i,j+1)) - H(\phi(i,j)))^2}}_{\triangleq l(i,j)}. \tag{B.2}$$

where $c_1(\phi)$ and $c_2(\phi)$ are the means of the function f (the image `im`) over the two regions $\{(i,j)|\, \phi(i,j) = 1\}$ and $\{(i,j)|\, \phi(i,j) = -1\}$. The reader would have recognized for (B.1) the Chan-Vese functional ([CV1]) which is the Mumford-Shah functional with two regions and for (B.2) the functional of Tsai, Yezzi and Willsky ([TYW]).

$$E(\phi) = \sum_{i,j} (\alpha - f(i,j)) H(\phi(i,j)) + \mu \sum_{i,j} \underbrace{\sqrt{(H(\phi(i+1,j)) - H(\phi(i,j)))^2 + (H(\phi(i,j+1)) - H(\phi(i,j)))^2}}_{\triangleq l(i,j)}. \tag{B.3}$$

which we will call threshold energy, since the case $\mu = 0$ corresponds to the thresholding of the function f (the image `im`) at the level α ($\phi = +1$ on the domain where $f(i,j) \leq \alpha$ and $\phi = -1$ on the complementary domain). When $\mu \neq 0$, the parameter μ acts as a regularizing parameter.
The energy minimization is performed using the classical ICM algorithm : the function ϕ is freezed on the pixels $(k,l) \neq (i,j)$, and we check the energy variation when the pixel (i,j) is changed from $\phi(i,j)$ to $-\phi(i,j)$; this procedure is applied for each pixel. In section 3.2.1.2, we have written the variation for the threshold energy.
For Chan-Vese energy, the variation for the first two terms is

$$\Delta E_1 + \Delta E_2 = (f(i,j) - c_2)^2 \frac{N_2}{N_2 + 1} - (f(i,j) - c_1)^2 \frac{N_1}{N_1 - 1}$$

if $\phi(i,j) = 1$, and where N_1 and N_2 are respectively the number of pixels in the regions $\phi = 1$ and $\phi = -1$, if $\phi(i,j) = -1$, just change (c_1, N_1) with (c_2, N_2) in the formula above.
For Tsai-Yezzi-Willsky energy, the variation for the first two terms is

$$\Delta E_{diff} = \left(\frac{f(i,j) - c_1}{N_1 - 1} + \frac{f(i,j) - c_2}{N_2 + 1}\right)^2 + 2\left(\frac{f(i,j) - c_1}{N_1 - 1} + \frac{f(i,j) - c_2}{N_2 + 1}\right)(c_2 - c_1)$$

if $\phi(i,j) = 1$, for the other case just change (c_1, N_1) with (c_2, N_2) in the formula above.

When selected, the option `-p` passes the precision to one of the two algorithms of Chan-Vese (equation B.1) and Tsai-Yezzi-Willsky (equation B.2). This means the algorithm stops when $|c_1^{n+1} - c_1^n| \leq$ `p` or $|c_2^{n+1} - c_2^n| \leq$ `p`, *i.e.* when the difference between the means is lower than the precision. When the option `-p` is not selected, the module performs the indicated number of sweeps `sweep`.
This algorithm is very fast (a few seconds) and usually at most 10 sweeps is enough to reach the convergence.
To apply this algorithm and to view the result upon an image, one can use the following script (two successive parameters α [the first one great and the second one lower] are used to denoise the segmentation).

```
#!/bin/sh

#1st parameter: threshold1, 2nd: image to be applied
#3rd: nb of sweeps, 4th: regularizing parameter,
#5th: image upon which the snake will be superposed ,
#6th: threshold2

binarize -t $1 $2 mask
fthre -n -m -1 -M 1 mask maskp1m1

icm 0 $3 $1 $4 $2 maskp1m1 icm_out
icm 0 $3 $6 $4 $2 icm_out icm_out

# just replace the two preceding lines to apply one
# of the other functionals

binarize -t 0 icm_out mask_out
fop -a 255 -m mask_out mask_out
ll_distance -l 255 mask_out $5.dist
emptypoly mask_out emptied
fmask -v 255 -c 255 $5.snake emptied $5


\rm mask maskp1m1 mask_out emptied
```

○Version 1.0

Last Modification date : Mon Mar 22 09 :50 :39 2004

○Author

F. Ranchin

Copyright (C) 1993-2004 *CMLA, ENS Cachan, 94235 Cachan cedex, France* - All rights reserved.

○Name

abjb Implementation of the piecewise constant variances active contours of Barlaud et al.

○Command Synopsis

abjb [-N *Nb_of_iterations*] [-p *p*] [-d *deltat*] [-n *niter*] [-D *delta*] [-e *eps*] [-m *mu*] [-t *t*] [-I *I*] [-l *lambda*] [-L *distance*] [-R *ref*] *fimage_polys fimage_ref abjb*

-N *Nb_of_iterations* : nb of iterations for moving (default :1)
-p *p* : precision to attain (no default value in order to)
-d *deltat* : time step of numerical scheme
-n *niter* : nb of iterations of reinitialization procedure
-D *delta* : time step of reinitialization procedure
-e *eps* : parameter used to approach absolute value
-m *mu* : scale parameter in the energy
-t *t* : threshold value, if set solves the PDE du/dt=| du| (t-im(x,y)+mu*curv(u))
-I *I* : another image : du/dt=delta(u)*(1/lambda*(alpha-phi(var)-phi'(var)*((ref-mean)^2-sigma^2))+mu*curv(u))
-l *lambda* : lambda parameter in this PDE
-L *distance* : distance to a reference mask
-R *ref* : reference image
fimage_polys : original polygon image (input)
fimage_ref : image on which the variances are computed
abjb : moved polygon image (output)

○Function Summary

Fimage abjb (a , im , I , Niter , p , deltat , niter , delta , eps , mu , lambda , t , distance , ref)
Fimage a , im , I , distance , ref ;
int *Niter , *niter ;
float *p ;
float *deltat , *delta , *eps , *mu , *lambda ;
float *t ;

○Description

This module implements the snakes algorithms described in [ABJB2]. Consider the energy

$$E(\Omega) = \phi(\sigma_\Omega^2)|\Omega| + \phi(\sigma_{D\setminus\Omega}^2)|D \setminus \Omega| + \mu\,\mathcal{H}^1(\partial\Omega).$$

We search the domain Ω in D for which this energy is minimal. We have shown in section 2.2.2.1 that this energy can also be expressed as a function of u, which is the distance function to the curve $\Gamma = \partial\Omega$

$$\begin{aligned} J_\alpha(u) = &\int_D \phi(\sigma^2)(\mathbf{x},u)H_\alpha(u)\,d\mathbf{x} + \int_D \phi(\sigma^2)(\mathbf{x},u)(1-H_\alpha(u))\,d\mathbf{x} \\ &+ \mu\int_D \delta_\alpha(u)|\nabla u|\,d\mathbf{x} \end{aligned}$$

where δ_α and H_α are C^1 or C^∞ approximations of the Dirac and Heaviside distributions. In our case, following the advise of Chan and Vese in [CV1], we have chosen the C^∞ approximations

$$\delta_\alpha(s) = \frac{2}{\pi}\frac{\alpha}{\alpha^2+s^2}$$

$$H_\alpha(s) = 1 + \frac{2}{\pi}\arctan(\frac{s}{\alpha}),$$

since this choice implies the use of the values $u(\mathbf{x})$ over the entire domain D (δ_α is never equal to 0).
The gradient descent of the energy u gives

$$\begin{aligned} \frac{\partial u}{\partial t} = \delta_\alpha(u)\Big(&\phi(\sigma_{D\setminus\Omega}^2) + \phi'(\sigma_{D\setminus\Omega}^2)((I-\mu_{D\setminus\Omega})^2 - \sigma_{D\setminus\Omega}^2) \\ &-\phi(\sigma_\Omega^2) - \phi'(\sigma_\Omega^2)((I-\mu_\Omega)^2-\sigma_\Omega^2) \\ &+\mu\,\mathrm{div}\left(\tfrac{\nabla u}{|\nabla u|}\right)\Big) \end{aligned}$$

This PDE is implemented using a semi-implicit scheme, that is to say the divergence term is approximated at the time $n+1$, and we use Romeny-Viergerer-Weickert AOS schemes (*cf.* section 2.4.2 for details).
If set, the option `-t` make the module apply the minimization of the another functional (the function f denotes the image `im`)

$$E(\Omega) = \alpha|\Omega| + \int_{D\setminus\Omega} f(\mathbf{x})\,d\mathbf{x} + \mu\mathcal{H}^1(\partial\Omega).$$

The transformation to a function of u is even more simple than in the first case, and the gradient descent gives

$$\frac{\partial u}{\partial t} \;=\; \delta_\alpha(u)\left(\alpha - f + \mu\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)\right).$$

When the initialisation you have chosen is quite far from an expected solution, the module may needs thousands of iterations and the algorithm will not be very fast. It would be a good idea to apply first an ICM algorithm before applying this module, as we have noticed it for the second functional in section 3.2.1.2.
Options `-n` and `-D` are free : they are not used by the module but are left to the user if he wants to add reinitialization to a signed distance function to the source code of the module. The `-I`, `-l`, `-R` and `-L` options are also free : they could be used for the implementation of the PDE indicated in the `-I` explanation. At last, the `-p` option is left to the user if he wants to add a stopping criterion to the source code.
Notice that the function *phi* is chosen as $\phi(s) = \log(1+s)$. For other functions, you need to modify the code. $\phi(s) = s$ corresponds to the Chan/Vese functional, see the next module which is exactly similar to this one for the options.
For a normal utilization of the module, you only have to use the `-N`, `-d` (you can choose it as 1. or even 5.), `-e` (usually 1. or between 1. and 2.) and `-m` options. Notice that the input needs a signed distance function to a set for a correct using of the module. To build such an image, you can use the following commands :

```
readpoly my_image my_curve
fillpoly -x 256 -y 256 my_curve my_curve
%for a 256*256 image, notice we replace the Fcurve my_curve
%by the binary Cimage of the domain it defines
ll_distance -l 255 my_curve my_dist
```

For transforming the output function to a mask, you simply have to binarize the function at the level 0

```
binarize -t 0 my_dist my_mask
```

And if you want to superimpose the snake upon the image, empty the mask and use fmask

```
emptypoly my_mask my_mask_emptied
fmask -v 255 -c 255 snake_image my_mask_emptied my_image
```

○Version 1.0

Last Modification date : Tue Mar 23 13 :38 :49 2004

○Author

F. Ranchin

Copyright (C) 1993-2004 *CMLA, ENS Cachan, 94235 Cachan cedex, France* - All rights reserved.

○Name

cv_ac Implementation of the Chan-Vese or Tsai-Yezzi-Willsky active contours using Weickert AOS scheme (separation of lines and columns treatment)

○Command Synopsis

cv_ac [-c *c*] [-N *Nb_of_iterations*] [-d *deltat*] [-n *niter*] [-D *delta*] [-e *eps*] [-m *mu*] *fimage_polys fimage_ref cv_ac*

-c *c* : Chan/Vese or Tsai/Yezzi/Willsky (c set) ?
-N *Nb_of_iterations* : nb of iterations for moving (default :1)
-d *deltat* : time step of numerical scheme
-n *niter* : nb of iterations of reinitialization procedure
-D *delta* : time step of reinitialization procedure
-e *eps* : parameter used to approach absolute value
-m *mu* : scale parameter in MS
fimage_polys : original polygon image (input)
fimage_ref :
cv_ac : moved polygon image (output)

○Function Summary

Fimage cv_ac (a , im , Niter , deltat , niter , delta , eps , mu , c)
Fimage a , im ;
int *Niter , *niter ;
float *deltat , *delta , *eps , *mu ;
char *c ;

○Description

This module is simply the preceding one with $\phi(s) = s$. The basic options are the same. The only different option is `-c`. When set, this make the module to achieve minimization of the Tsai-Yezzi-Willsky functional

$$E(\Omega) = -\frac{1}{2}|\mu(\Omega) - \mu(D \setminus \Omega)|^2 + \mu\mathcal{H}^1(\partial\Omega).$$

The minimization can be achieved using a gradient descent on the functional with the distance function u

$$\frac{\partial u}{\partial t} = \delta_\alpha(u)(\mu(\Omega) - \mu(D \setminus \Omega)) \left(\frac{I - \mu(\Omega)}{|\Omega|} + \frac{I - \mu(D \setminus \Omega)}{|D \setminus \Omega|} \right).$$

The implementation uses the same parameters than for Chan-Vese functional.

○Version 1.0

Last Modification date : Tue Mar 23 14 :13 :50 2004

○Author

F. Ranchin

Copyright (C) 1993-2004 *CMLA, ENS Cachan, 94235 Cachan cedex, France* - All rights reserved.

○Name

gac Implementation of the geodesic/geometric/MCM active contours using Romeny-Viergerer-Weickert AOS scheme (separation of lines and columns treatment)

○Command Synopsis

gac [-c *c*] [-N *Nb_of_iterations*] [-d *deltat*] [-l *lambda*] [-s *sigma*] [-k *k*] *fimage_polys fimage_ref gac_new*

-c *c* : what type of active contours (geometric, geodesic or MCM)
-N *Nb_of_iterations* : nb of iterations for moving (default : 100)
-d *deltat* : time step of numerical scheme
-l *lambda* : scale parameter in the edge detector function g=1/(1+| nabla I|ˆ2/lambdaˆ2) (default : 1)
-s *sigma* : standard deviation of the gaussian used for filtering
-k *k* : weight of the balloon force
fimage_polys : original polygon image (input)
fimage_ref :
gac_new : moved polygon image (output)

○Function Summary

Fimage gac (c , Niter , deltat , lambda , sigma , k , a , DI)
int *c ;
int *Niter ;
float *deltat , *lambda , *sigma , *k ;
Fimage a , DI ;

○Description

This module implements the geodesic/geometric/MCM (option `-c` set to 1, 2 and 0 respectively) active contours algorithm ([CKS]) via level sets and with an AOS numerical scheme (see [RVW]). The geodesic active contours have been introduced by Caselles, Kimmel and Sapiro. The aim is to find a sufficiently contrasted curve Γ in the image, *i.e.* a curve which minimizes

$$\int_\Gamma \frac{1}{1+|G_\sigma * \nabla I|^2(\mathbf{x})}\, d\mathcal{H}^1(\mathbf{x}) + \nu \int_\Omega \frac{1}{1+|G_\sigma * \nabla I|^2(\mathbf{x})}\, d\mathbf{x}$$

where Ω is the set enclosed by the curve Γ. The gradient descent of this energy gives

$$\frac{\partial \Gamma}{\partial t} = -\left(g\kappa + < \nabla g, \mathbf{n} > + \nu\right)\mathbf{n}$$

where $g(\mathbf{x}) = \frac{1}{1+|G_\sigma * \nabla I|^2(\mathbf{x})}$, κ denotes the curvature of Γ and $\mathbf{n}$ is the outward normal. Using level sets (see section 2.1.4.2 for explanation), we can rewrite this PDE into a PDE on the signed distance function u to the curve

$$\frac{\partial u}{\partial t} = |\nabla u| \left(\mathrm{div}\left(g\frac{\nabla u}{|\nabla u|}\right) + \nu\right).$$

The geometric snakes have been introduced by Caselles, Catté, Coll and Dibos in [CCCD]. This is a PDE model which is quite similar to the previous one since it is given by

$$\frac{\partial u}{\partial t} = g|\nabla u| \left(\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) + \nu\right).$$

For the two models, ν is set by the option `-k`. Be careful to set a negative value, since the convention we have chosen for the signed distance function is u positive inside the curve.
Whatever any other option used, `-s` takes the value σ used for the convolution of ∇I.

○Version 1.0

Last Modification date : Mon Jul 5 10 :59 :30 2004

○Author

F. Ranchin

Copyright (C) 1993-2004 *CMLA, ENS Cachan, 94235 Cachan cedex, France* - All rights reserved.

○Name

ws_flow Weickert and Schnoerr optical flow computation

○Command Synopsis

ws_flow [-p *precision*] [-n *n*] [-t *tau*] [-l *lambda*] [-E *eps*] [-A *alpha*] [-R *norm_movie*] [-N *angle_movie*] [-U *fmoviex_out*] [-V *fmoviey_out*] [-D *discontinuity*] *fmovie*

-p *precision* : stop the algorithm when the absolute residue become inferior to [percentage] of the first residue
-n *n* : number of iterations for numerical scheme
-t *tau* : time-step for each treatment of pde(sigma)
-l *lambda* : contrast parameter
-E *eps* : epsilon parameter (just necessary for theory, you can choose it as weak as possible)
-A *alpha* : parameter which represents weight of the divergence in pde
-R *norm_movie* : computes optical flow norm if selected
-N *angle_movie* : computes optical flow direction if selected
-U *fmoviex_out* : computes optical flow first cartesian component if selected
-V *fmoviey_out* : computes optical flow second cartesian component if selected
-D *discontinuity* : flow discontinuity (in spatio-temporal domain)
fmovie : Input movie

○Function Summary

void ws_flow (percent , n , tau , lambda , eps , alpha , norm , dir , movie , wsU , wsV , discontinuity)
float *percent ;
int *n ;
float *tau , *lambda , *eps , *alpha ;
Fmovie norm , dir ;
Fmovie movie , wsU , wsV , discontinuity ;

○Description

ws_flow is an implementation of the Weickert and Schnörr optical flow computation by a semi-implicit scheme (in comparison of other schemes, it is quasi-

explicit). In [WS], they consider a functional where the gradients of the two components of the flow are 3D-gradients, *i.e.* $\nabla^3 f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial t})^T$ and they do not separate the 3D-gradients of the two components (σ_1, σ_2) of the velocity

$$E(\sigma) = \int_{\Omega\times[0,T]} |\nabla u \cdot \sigma + u_t|^2 \, dx\, dy\, dt + \alpha \int_{\Omega\times[0,T]} \Psi(|\nabla^3\sigma_1|^2 + |\nabla^3\sigma_2|^2)\, dx\, dy\, dt$$

where u is the gray level at pixel (x, y) and time t, $\lambda > 0$ and $\Psi(s^2) = \epsilon s^2 + (1 - \epsilon)\lambda^2\sqrt{1 + \frac{s^2}{\lambda^2}}$. ($\epsilon$ is required only for proving well-posedness and can be chosen as weak as possible, e.g. $\epsilon = 10^{-6}$). The steepest descent equations are

$$\begin{aligned} \frac{\partial\sigma_1}{\partial t} &= \mathrm{div}^3(\Psi'(|\nabla^3\sigma_1|^2 + |\nabla^3\sigma_2|^2)\nabla^3\sigma_1) - \tfrac{1}{\alpha}\, u_x(\nabla u \cdot \sigma + u_t) \\ \frac{\partial\sigma_2}{\partial t} &= \mathrm{div}^3(\Psi'(|\nabla^3\sigma_1|^2 + |\nabla^3\sigma_2|^2)\nabla^3\sigma_2) - \tfrac{1}{\alpha}\, u_y(\nabla u \cdot \sigma + u_t) \end{aligned}$$

The semi-implicit scheme consists in approximating $\frac{\partial\sigma_1}{\partial t}$ by an Euler-forward scheme $\frac{\sigma_1^{n+1} - \sigma_1^n}{\delta t}$, $\mathrm{div}^3(\Psi'(|\nabla^3\sigma_1|^2 + |\nabla^3\sigma_2|^2)\nabla^3\sigma_1)$ at time n by

$$\sum_{(i,j)\in\mathcal{N}_6(x,y)} w(\sigma_1^n(i,j))\sigma_1^n(i,j)$$

($\mathcal{N}_6(x, y)$ is the 6-neighbourhood of pixel (x, y) : 4 neighbours in space+2 neighbours in time) ; $\frac{1}{\alpha}u_x(\nabla u \cdot \sigma + u_t)$ by

$$\frac{1}{\alpha}\, u_x(u_x\sigma_1^{n+1} + u_y\sigma_2^n + u_t)$$

and the same way for the second equation.
The values $w(\sigma_1^n(i, j))$ come from the Malik and Perona discretization of the divergence (approximation of the derivatives by a centered scheme at semi-nodes)

$$w(\sigma_1^n(i,j))\sigma_1^n(i,j) = \frac{\Psi'(i,j,n) + \Psi'(x,y,n)}{2}(\sigma_1^n(i,j) - \sigma_1^n(x,y))$$

where $\Psi'(i, j, n)$ approximates $\Psi'(|\nabla^3\sigma_1^n|^2 + |\nabla^3\sigma_2^n|^2)(i, j)$.
This semi-implicit scheme is a median way between a completely explicit scheme and AOS schemes (see [RVW]).
The iterations are performed until `n` is reached, unless the precision `percent` is attained. There is no default value for this last parameter, in order to do all the iterations until `n` in case the user does not want to use the option `-p`, in the opposite case a value must be chosen.

○Version 1.0

Last Modification date : Mon Mar 31 19 :28 :39 2003

○Author

Florent Ranchin

Copyright (C) 1993-2003 *CMLA, ENS Cachan, 94235 Cachan cedex, France* - All rights reserved.

Références

[AKP] **T. Aach, A. Kaup and R. Mester**, *Statistical model-based change detection in moving video*, Signal Processing **31 2** (1993), 165-180;

[Al] **L. Alvarez**, *Images and PDE's*, 12$^{\text{th}}$ International Conference on Analysis and Optimization of Systems, Images, Wavelets and PDE's, Springer (M.-O. Berger, R. Deriche, I. Herlin, J. Jaffré and J.-M. Morel, eds.) (1996), 3-14;

[AGLM] **L. Alvarez, F. Guichard, P.-L. Lions and J.-M. Morel**, *Axioms and fundamental equations of image processing*, Arch. for Rat. Mech. **16, IX** (1992), 200-257;

[AWS] **L. Alvarez, J. Weickert, J. Sánchez**, *Reliable Estimation of Dense Optical Flow Fields with Large Displacements*, International Journal of Computer Vision **39** (2000), 41-56;

[ABGJB] **G. Aubert, M. Barlaud, M. Gastaud and S. Jehan-Besson**, *Region-Based Active Contours Using Geometrical and Statistical Features for Image Segmentation*, International Conference on Image Processing Proceedings (2003), Barcelona, Spain;

[ABFJB] **G. Aubert, M. Barlaud, O. Faugeras and S. Jehan-Besson**, *Image Segmentation Using Active Contours : Calculus of Variations or Shape Gradients ?*, SIAM Journal of Applied Mathematics **63 6** (2003), 2128-2154;

[ABJB1] **S. Jehan-Besson, M. Barlaud and G. Aubert**, *DREAM^{2}S : Deformable Regions driven by an Eulerian Accurate Minimization Method for image and video segmentation*, International Journal of Computer Vision **53 1** (2003), 45-70;

[ABJB2] **S. Jehan-Besson, G. Aubert, M. Barlaud**, *Video Object Segmentation Using Eulerian Region-Based Active Contours*, International Conference in Computer Vision (2001), Vancouver, Canada;

[ABJB3] **S. Jehan-Besson, G. Aubert, M. Barlaud**, *Region-based active contours for video object segmentation with camera compensation*, International Conference on Image Processing Proceedings (2001), Thessaloniki, Greece;

[ADK1] **G. Aubert, R. Deriche and P. Kornprobst**, *A mathematical study of regularized optical flow problem in space BV*, SIAM. Journal on Mathematical analysis **30** (1999), 1282-1308;

[ADK2] **G. Aubert, R. Deriche and P. Kornprobst**, *Image Sequence Analysis via Partial Differential Equations*, Journal of Mathematical Imaging and Vision **11 1** (1999), 5-26;

[AK] **G. Aubert and P. Kornprobst**, Mathematical problems in image processing : Partial Differential Equations and the Calculus of Variations, Springer-Verlag, Applied Mathematics and Science 147, 2001;

[BP] **M. Barlaud and F. Precioso**, *B-spline active contours for fast video segmentation*, International Conference on Image Processing Proceedings (2001), Thessaloniki, Greece;

[Ber] **C. Bernard**, *Fast Optic Flow with Wavelets*, Proceedings of the Wavelet and Applications Workshop (1998);

[BSR1] **M. Bertalmío, G. Sapiro and G. Randall**, *Morphing Active Contours*, Proceedings IEEE-International Conference on Image Processing (1998), Chicago;

[BSR2] **M. Bertalmío, G. Sapiro and G. Randall**, *Region tracking on level sets methods*, IEEE transactions on Medical Imaging **18 n° 5** (1999), 448-451;

[Be] **J. Besag**, *On the statistical analysis of dirty pictures*, Journal of the Royal Statistical Society **B-48** (1986), 259-302;

[BGW] **J. Bigün, G. H. Granlund and J. Wiklund**, *Multidimensional orientation estimation with applications to texture analysis and optical flow*, IEEE Transactions on Pattern Analysis and Machine Intelligence **n° 8** (1991), 775-790;

[BZ] **A. Blake and A. Zisserman**, VISUAL RECONSTRUCTION, Cambridge, Mass. : MIT Press, 1987;

[BW] **T. Brox and J. Weickert**, *Nonlinear matrix diffusion for optic flow estimation*, Pattern Recognition, Lecture Notes in Computer Science Springer (L. Van Gool (Ed.), eds.) **2449** (2002), 446-453;

[BWFKS] **A. Bruhn, J. Weickert, C. Feddern, T. Kohlberger and C. Schnörr**, *Variational optic flow computation in real-time*, IEEE Transactions in Image Processing (to appear);

[KB] **A. Bruckstein and R. Kimmel**, *On edge detection, edge integration and geometric active contours*, ISMM 02 Proceedings (2002);

[CCCD] **V. Caselles, F. Catté, B. Coll and F. Dibos**, *A geometric model for active contours in image processing*, Numerische Mathematik **66** (1993), 1-31;

[CC] **V. Caselles and B. Coll**, *Snakes in movement*, SIAM Journal of Numerical Analysis **33(6)** (1996), 2445-2456;

[CKS] **V. Caselles, R. Kimmel and G. Sapiro**, *Geodesic Active Contours*, International Conference on Computer Vision (1995), 694-699;

[CDK] **F. Catté, F. Dibos and G. Koepfler**, *A Morphological Scheme for Mean Curvature Motion and Applications to Anisotropic Diffusion and Motion of Level Sets*, SIAM Journal on Numerical Analysis **32 n° 6** (1995), 1895–1909;

[Cham] **A. Chambolle**, *Image segmentation by variational methods : Mumford and Shah functional and the discrete approximations*, SIAM Journal of Applied Mathematics **55 n° 3** (1995), 827-863;

[CV1] **T. F. Chan and L. A. Vese**, *Active contours without edges*, IEEE Transactions on Image Processing **10(2)** (2001), 266-277;

[CV2] **T. F. Chan and L. A. Vese**, *A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model*, International Journal of Computer Vision **50(3)** (2002), 271-293;

[CV3] **T. F. Chan and L. A. Vese**, *An Efficient Variational Multiphase Motion for the Mumford-Shah Segmentation Model*, Proceedings of the 34'th Asilomar Conference on Signals, Systems, and Computers **1** (2000), 490-494;

[CBFAB] **P. Charbonnier, L. Blanc-Féraud, G. Aubert and M. Barlaud**, *Deterministic edge preserving regularization in computed imaging*, IEEE Transactions on Image Processing **6 2** (1997), 298-311;

[CheV] **Y. Chen and B. Vemuri**, *Joint Image registration and segmentation*, Geometric Level Set Methods in Imaging, Vision and Graphics, Springer Verlag (Stanley Osher and Nikos Paragios, eds.) (2003);

[Co] **L. Cohen**, *On active contour models and balloons*, Computer Vision Graphics and Image Processing : Image understanding **53** (1991), 211-218;

[Co2] **L. Cohen**, *Avoiding Local Minima for Deformable Curves in Image Analysis*, Curves and Surfaces with Applications in CAGD (1997);

[CBA] **L. Cohen, E. Bardinet, and N. Ayache**, *Surface reconstruction using active contour models*, Proceedings SPIE 93 Conference on Geometric Methods in Computer Vision (1993);

[CEGLM] **T. Cohignac, F. Eve, F. Guichard, C. Lopez and J.-M. Morel**, *Numerical Analysis of the Fundamental Equation of Image Processing*, Les Cahiers du CEREMADE n° 9254;

[CS] **D. Cremers and C. Schnörr**, *Motion Competition : Variational Integration of Motion Segmentation and Shape Regularization*, Proceedings of the German Conf. on Pattern Recognition (DAGM), Springer LNCS (L. Van Gool, eds.) **2449** (2002), 472-480;

[Da] **P. Davies**, *Aspects of robust linear regression*, The Annals of Statistics **21** (1993), 1843-1889;

[DZ] **M. Delfour and J.-P. Zolésio**, SHAPES AND GEOMETRIES, SIAM, Philadelphia, PA, Advances in Design and Control, 2001;

[DP1] **R. Deriche and N. Paragios**, *Geodesic Active Regions for Motion Estimation and Tracking*, INRIA Research Report n° 3631 (1999);

[DP2] **R. Deriche and N. Paragios**, *Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects*, IEEE Transactions on Pattern Analysis and Machine Intelligence (1999);

[DMM] **A. Desolneux, L. Moisan and J.-M. Morel**, *Variational snakes theory*, Geometric Level Set Methods in Imaging, Vision and Graphics, Springer Verlag (Stanley Osher and Nikos Paragios, eds.) (2003);

[Di] **F. Dibos**, *Du groupe projectif au groupe des recalages : une nouvelle modélisation*, C. R. de l'Académie des Sciences **332(9)** (2001);

[ES] **L. C. Evans and J. Spruck**, *Motion of level sets by mean curvature*, Journal of Differential Geometry **33** (1991), 635-681;

[FG] **J. Gomes and O. Faugeras**, *Reconciling Distance Functions and Level Sets*, INRIA report n° 3666 (1999);

[Ge] **S. Geman and D. Geman**, *Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images*, IEEE Transactions on Pattern Analysis and Machine Intelligence **6** (1993), 721-741;

[GM] **F. Guichard and J.-M. Morel**, IMAGE ANALYSIS AND PDE'S, to appear[1], 2001;

[HR] **M. Hintermüller and W. Ring**, *A second order shape optimization approach for image segmentation*, Report 237, Special Research Center on Optimization and Control, University of Graz, Austria; to appear in SIAM J. Appl. Math. (2002);

[HS] **B.K.P. Horn and B. Schunck**, *Determinating Optical Flow. Artificial intelligence*, Artificial intelligence (1981);

[Hu] **P. Huber**, ROBUST STATISTICS, Wiley, 1981;

[Ka] **G. Kanisza**, GRAMMAIRE DU VOIR, Diderot, 1996;

[KWT] **M. Kass, A. Witkin and D. Terzopoulos**, *Snakes: Active contour models*, International Journal of Computer Vision **1** (1988), 321-332;

[KLM] **G. Koepfler, C. Lopez and J.-M. Morel**, *A Multiscale Algorithm For Image Segmentation By Variational Method*, SIAM Journal on Numerical Analysis **31(1)** (1994), 282–299;

[Ko] **P. Kornprobst**, *Contribution à la restauration d'images et à l'analyse de séquences : Approches Variationnelles et Solutions de Viscosité*, Ph.D. thesis (1998);

[KWSR] **G. Kühne, J. Weickert, O. Schuster, S. Richter**, *A tensor-driven active contour model for moving object segmentation*, IEEE Proceedings on International Conference on Image Processing **2** (2001), 73-76;

[LKM] **F. Lauze, P. Kornprobst, E. Mémin**, *A Coarse To Fine Multiscale Approach For Linear Least Squares Optical Flow Estimation*, BMCV Proceedings (2004);

[LHC] **B. Leroy, I. Herlin and L. D. Cohen**, *Multi-Resolution algorithms for Active Contour Models*, ICAOS 96, Images, Wavelets and PDEs, Springer Verlag (M.-O. Berger, R. Deriche, I Herlin, J. Jaffré and J.-M. Morel, eds.) (1996);

[1] available online at `www.ipam.ucla.edu/publications/gbm2001/gbmtut_jmorel.pdf` (all rights reserved).

[LK] **B. Lucas and T. Kanade**, *An iterative image registration technique with an application to stereo vision*, International Joint Conference on Artificial Intelligence (1981), 674-679;

[MSV] **R. Malladi, J. Sethian and B. Vemuri**, *Shape modeling with front propagation: a level set approach*, IEEE Transactions on Pattern Analysis and Machine Intelligence **17** (1995), 158-175;

[MMA] **A.-R. Mansouri and A. Mitiche and M. Aron**, *PDE based Tracking Without Motion Computation by Joint Space-Time Segmentation*, IEEE International Conference on Image Processing (2003), III:113-116;

[MCZM] **F. Moschetti, G. Covitto, F. Ziliani and A. Mecocci**, *Automatic object extraction and bitrate allocation for second generation video coding*, Proceedings of ICME 02 (2002);

[MP] **E. Mémin and P. Perez**, *A Multigrid approach for hierarchical motion estimation*, Proceedings of the 6th International Journal of Computer Vision, IEEE Computer Society Press (1998), 933-938;

[MS1] **D. Mumford and J. Shah**, *Boundary detection by minimizing functionals*, Proceedings of the International Conference on Computer Vision and Pattern Recognition, IEEE, San Francisco , CA (1985), 22-26;

[MS2] **D. Mumford and J. Shah**, *Optimal approximations by piecewise smooth functions and associated variational problems*, Communications on Pure and Applied Mathematics **42** (1989), 577-684;

[NE] **H.-H. Nagel and W. Enkelmann**, *An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences*, IEEE Trans. Pattern Anal. Mach. Intel. **8** (1986), 565-593;

[NM] **H.-H. Nagel and M. Middendorf**, *Estimation and Interpretation of Discontinuities in Optical Flow Fields*, ICCV Proceedings (2001), 178-183;

[OB] **J.-M. Odobez and P. Bouthemy**, *Robust multiresolution estimation of parametric motion models*, Journal of Visual Communication and Image Representation **6** (1995), 348-365;

[OS] **S. Osher and J. Sethian**, *Fronts propagating with curvature dependent speed: algorithms based on the Hamilton-Jacobi formulation*, Journal of Computational Physics **79** (1990), 12-49;

[PT] **N. Paragios and G. Tziritas**, *Detection and location of moving objects using deterministic relaxation algorithms*, Proceedings of IEEE CVPR (1996);

[PM] **P. Perona and J. Malik**, *Scale space and edge detection using anisotropic difusion*, Proc. IEEE Computer Soc. Workshop on Computer Vision (1987);

[RVW] **B. M. ter Haar Romeny, M. A. Viergerer and J. Weickert**, *Efficient and Reliable Schemes for Nonlinear Diffusion Filtering*, IEEE Trans. Image Proc. (1998);

[Ro] **R. Ronfard**, *Region-based strategies for active contour models*, International Conference in Computer Vision Proceedings **13 n° 2** (1994), 229-251;

[SBFAZ] **C. Samson, L. Blanc-Féraud, G. Aubert, and J. Zérubia**, *A variational model for image classification and restoration*, IEEE Transactions on Pattern Analysis and Machine Intelligence **22(5)** (2000), 460-472;

[Sa] **O. Sanchez**, *Désoccultation de séquences vidéo : aspects mathématiques et numériques*, Ph.D. thesis (2002);

[SD] **O. Sanchez and F. Dibos**, *Displacement Following of Hidden Objects in a Video Sequence*, International Journal of Computer Vision **57 n° 2** (2004), 91-105;

[SZ] **J. Sokolowski and J.-P. Zolésio**, *Introduction to Shape Optimization. Shape sensitivity analysis*, Springer Ser. Comput. Math., Springer-Verlag (1992);

[SC] **B. Song and T. F. Chan**, *A Fast Algorithm for Level Set Based Optimization*, UCLA Cam Report **02-68** (2002);

[TYW] **A. Tsai, A. Yezzi and A. Willsky**, *A Fully Global Approach to Image Segmentation via Coupled Curve Evolution Equations*, Journal of Visual Communication and Image Representation **13 n° 1/2** (2002), 195-216;

[Vi] **L. Vincent**, *Grayscale area openings and closings, their efficient implementation and applications*, Proceedings of the First Workshop on Mathematical Morphology and its Applications to Signal Processing **A** (1994), 109-114;

[We] **J. Weickert**, Anisotropic Diffusion in Image Processing, ECMI, Teubner Verlag, 1998;

[WK] **J. Weickert and G. Kühne**, *Fast Methods for Implicit Active Contour Models*, Geometric Level Set Methods in Imaging, Vision and Graphics, Springer Verlag (Stanley Osher and Nikos Paragios, eds.) (2003);

[WS] **J. Weickert and C. Schnörr**, *Variational Optic flow Computation with a spatio-Temporal Smoothness Constraint*, Journal of Mathematical Imaging and Vision, Kluwer Academic Publishers **14** (2001), 245-255;

[ZY] **S.-C. Zhu and A. Yuille**, *Region competition : unifying snakes, region growing, and Bayes/MDL for multiband image segmentation*, IEEE Transactions on Pattern Analysis and Machine Intelligence **18** (1996), 884-900.