

# THESE

pour l'obtention du titre de  
**DOCTEUR EN SCIENCES**  
spécialité : Mathématiques Appliquées  
arrêté du 23 novembre 1988

**Sujet :**

TRAITEMENT NUMERIQUE D'IMAGES ET DE FILMS :  
EQUATIONS AUX DERIVEES PARTIELLES  
PRESERVANT FORME ET RELIEF.

**Candidat :** Lionel MOISAN

**Directeur de recherche :** Jean-Michel MOREL

**Jury :**

Président :	Monsieur Yves MEYER
Rapporteur :	Monsieur Guy BARLES
Rapporteur :	Monsieur Vicent CASELLES
Examineur :	Monsieur Vladimir ARNOLD
Examineur :	Monsieur Robert AZENCOTT
Examineur :	Monsieur Jean-Michel MOREL
Examineur :	Monsieur Leonid YAROSLAVSKY

**juin 1997**

*L'université n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses : ces opinions doivent être considérées comme propres à leurs auteurs.*



*A mes parents,  
à Claire.*



*Monsieur Yves Meyer a accepté de présider ce jury. Je voudrais lui exprimer ici ma très vive gratitude pour cette attention portée à mon travail et pour l'honneur qu'il me fait ainsi.*

*Je tiens à remercier particulièrement Monsieur Vicent Caselles pour ses conseils et ses encouragements qui m'ont été extrêmement précieux : nos discussions ont toujours été pour moi une source stimulante de réflexion. Pour avoir rapporté cette thèse et pour toute son aide, qu'il soit assuré de ma très profonde reconnaissance.*

*A Monsieur Guy Barles, je souhaite également dire ma sincère gratitude, pour avoir accepté aussi la charge de rapporteur, et pour la grande disponibilité dont il a fait preuve à cette occasion.*

*Monsieur Robert Azencott, qui m'a accueilli comme caïman à l'Ecole Normale Supérieure de Cachan avec Monsieur Jean-Michel Ghidaglia, a accepté de se joindre à ce jury. Qu'il me permette de lui dire ici ma reconnaissance. Mes remerciements vont aussi à Monsieur Yaroslavsky, qui participe à ce jury, et à Monsieur Vladimir Arnold, qui me fait également l'honneur de s'y joindre, malgré un emploi du temps que je sais très chargé.*

*Pour une collaboration stimulante, je voudrais remercier toute l'équipe de traitement d'images du CEREMADE, et particulièrement Francine Catté, Thierry Cohignac, Françoise Dibos, Jacques Froment, Georges Koepfler et Christian Lopez. Merci également aux membres du CMLA et du département de mathématiques de l'ENS Cachan pour leur accueil sympathique. Pour m'avoir fait découvrir le traitement d'images en milieu industriel, merci à Jean-Paul Boujot et à la CISI, ainsi qu'à Pierre-Louis Lions pour les discussions très enrichissantes que nous avons eues à ces occasions. Je remercie également Thierry Cohignac et le Ministère de la Défense, grâce à qui j'ai pu effectuer mon service militaire dans les meilleures conditions d'un laboratoire de recherche.*

*Sans Michel Vanbreugel, cette thèse n'aurait pas résisté à la loi de Murphy. Pour avoir su maîtriser un parc informatique toujours grandissant, et pour m'avoir ainsi permis de travailler dans les meilleures conditions, merci à lui. Merci également à Lucette Jean Pierre, Josette Lévy, Magali Marc-Dibildos et Irène Mazzella pour l'aide qu'elles m'ont apportée dans toutes les démarches administratives.*

*Je souhaite également exprimer ma profonde sympathie à Jean-Pierre d'Alès, Mohamed Bouguecha, Kaouther Boussema, Frédéric Cao, Rafia Elkalay, Orlando Fernandez, Yann Gousseau, Frédéric Guichard, David Khuat Duy, Simon Masnou, Pascal Monasse, Jacques-Olivier Moussafir, Denis Pasquignon et Fatma Trigui, pour des discussions toujours passionnantes, pour leur soutien moral, et pour la bonne humeur et l'atmosphère de recherche motivante qui règnent dans les bureaux du CEREMADE. Merci aussi à mes amis de L'Ecole Normale Supérieure pour le recul et l'ouverture d'esprit qu'ils ont su m'apporter.*

*Enfin, je ne saurais comment remercier Monsieur Jean-Michel Morel, qui a dirigé cette thèse, m'a guidé et soutenu tout au long de ces années de recherche. Son attention constante à mon travail, les discussions extrêmement stimulantes par lesquelles il m'a fait découvrir le monde passionnant de la recherche, ont été pour moi une aide irremplaçable. Pour tout ce qu'il m'a ainsi apporté, je veux lui témoigner ici ma très profonde et sincère reconnaissance.*



## Résumé

La reconnaissance de formes planes partiellement masquées ne peut se faire que localement en calculant des points caractéristiques (extremas de courbure, points d'inflexions,...), et ce calcul requiert un procédé de lissage des formes. Si l'on veut effectuer cette reconnaissance modulo toutes les déformations affines du plan, alors ce procédé est unique : c'est le *scale space affine*, découvert en 1993, qui peut être décrit par une équation d'évolution. Dans la première partie de cette thèse, nous montrons comment résoudre cette équation numériquement avec précision, en itérant un opérateur continu, géométrique, global et exactement calculable. Des propriétés de consistance forte et de convergence sont établies et validées par de nombreuses expériences numériques. Ce procédé offre des performances bien supérieures aux schémas classiques aux différences finies, qui ne peuvent vérifier rigoureusement l'invariance affine et le principe d'inclusion.

Dans une deuxième partie, nous étudions l'un des problèmes fondamentaux de la robotique, la reconstruction du relief à partir d'une séquence d'images. Il s'avère que lorsque le mouvement de l'observateur peut être déterminé, il n'existe fondamentalement qu'une seule manière de filtrer la séquence d'images tout en préservant le relief sous-jacent. Ce filtrage, obtenu grâce à une démarche axiomatique, se formule par une équation aux dérivées partielles non linéaire du second ordre, parabolique dégénérée, qui présente une singularité très forte inhérente au problème de reconstruction. Nous établissons des résultats d'existence et d'unicité pour cette équation, puis mettons en évidence certaines propriétés mathématiques qui se prêtent facilement à une interprétation physique. Enfin, nous décrivons un schéma numérique adapté, et réalisons des expériences qui montrent que ce filtrage, par la cohérence qu'il induit, ramène le procédé de reconstruction à un calcul élémentaire et fiable.

## Abstract

The recognition of partially occluded planar shapes necessarily involves a local computation of characteristic points (curvature extrema, inflexion points,...), and this computation requires a shape smoothing process. If the recognition is considered up to all affine transformations of the plane, then this process is unique : this is the *affine scale space*, discovered in 1993, which can be described by an evolution equation. In the first part of this study, we show how this equation can be solved numerically with a high accuracy, by iterating a continuous, geometric and global operator which can be exactly computed. Full consistency and convergence results are provided, as well as conclusive numerical experiments. This method goes beyond classical finite differences schemes that never manage to satisfy rigorously the affine invariance and the inclusion principle.

In a second part, we study a fundamental problem of robotics : the depth recovery from a sequence of images. We prove that when the camera motion can be controlled, there fundamentally exists only one way to process the image sequence and preserve the underlying depth in the same time. This process, obtained from an axiomatic formulation, can be described by a non-linear second-order degenerate parabolic partial differential equation, which presents a strong singularity inherent to the depth recovery problem. We establish existence and uniqueness results for this equation, and we highlight several properties which can be easily interpreted from a physical point of view. Last, we describe a numerical scheme, and show on experiments how the filtering process, thanks to the coherence it induces, brings back the depth recovery to an elementary and robust computation.





# Contents

<b>1</b>	<b>Introduction</b>	<b>13</b>
1.1	Problèmes et Enjeux du traitement d'images . . . . .	13
1.2	Analyse des images : la notion de <i>scale space</i> . . . . .	14
1.3	Plan de l'exposé. . . . .	16
<b>I</b>	<b>A strongly consistent geometrical scheme for the Affine Scale Space</b>	<b>21</b>
<b>2</b>	<b>The Affine Scale Space</b>	<b>23</b>
2.1	Image analysis and scale spaces . . . . .	23
2.2	Definition . . . . .	24
2.2.1	Image formulation . . . . .	24
2.2.2	Geometric formulation . . . . .	26
2.2.3	Applications . . . . .	28
2.3	Numerical schemes for the Affine Scale Space . . . . .	30
2.3.1	Definitions . . . . .	30
2.3.2	The Osher-Sethian's method . . . . .	30
2.3.3	State of the art . . . . .	30
2.3.4	Point evolution schemes . . . . .	32
2.4	A fully consistent scheme . . . . .	32
<b>3</b>	<b>Affine erosion of curves and sets</b>	<b>35</b>
3.1	Preliminaries . . . . .	35
3.2	Affine erosion of sets . . . . .	38
3.2.1	Definition . . . . .	38
3.2.2	Example . . . . .	39
3.2.3	Topological structure . . . . .	40
3.2.4	Affine dilation . . . . .	42
3.2.5	Basic properties of the affine erosion . . . . .	42
3.3	Affine erosion of convex curves . . . . .	45
3.3.1	Basic statements . . . . .	45
3.3.2	The middle point property . . . . .	46
3.3.3	Regular scales . . . . .	52
3.3.4	Consistency . . . . .	55
3.4	Affine erosion of non convex curves . . . . .	60
3.4.1	Structure . . . . .	60
3.4.2	Inflexion points . . . . .	62

3.4.3	Consistency . . . . .	64
3.4.4	Other possible definitions of the affine erosion . . . . .	65
<b>4</b>	<b>Comparison between affine erosion and scale space</b>	<b>67</b>
4.1	Affine scale space of curves . . . . .	67
4.2	Affine erosion and scale space of an ellipse . . . . .	68
4.2.1	Affine erosion . . . . .	68
4.2.2	Affine scale space . . . . .	70
4.3	Affine erosion and scale space of a hyperbola . . . . .	72
4.3.1	Affine erosion . . . . .	72
4.3.2	Affine scale space . . . . .	74
4.4	Affine erosion and scale space of a parabola . . . . .	75
4.5	Affine erosion of a triangle . . . . .	77
<b>5</b>	<b>Affine erosion of grey-level images</b>	<b>81</b>
5.1	Morphological principles . . . . .	81
5.2	From sets to images . . . . .	82
5.3	Affine erosion of grey level images . . . . .	84
5.4	Comparison with the inf-sup operators . . . . .	89
5.5	Asymptotic behaviour of the affine erosion . . . . .	94
5.5.1	A local comparison principle . . . . .	94
5.5.2	Consistency . . . . .	97
5.6	Using Matheron's Theorem . . . . .	100
5.7	Convergence . . . . .	103
<b>6</b>	<b>Numerical scheme</b>	<b>105</b>
6.1	Affine erosion of a polygon . . . . .	106
6.1.1	Regular convex case . . . . .	106
6.1.2	Non regular convex case (removing ghosts parts) . . . . .	108
6.1.3	General case (non convex polygons) . . . . .	109
6.2	Algorithm . . . . .	113
6.3	Affine subsampling and iteration . . . . .	116
6.4	A simplified algorithm . . . . .	117
6.4.1	Pseudo affine erosion . . . . .	117
6.4.2	Algorithm . . . . .	119
<b>7</b>	<b>Experiments</b>	<b>121</b>
7.1	Affine erosions . . . . .	121
7.1.1	Discretization . . . . .	128
7.1.2	Affine Invariance . . . . .	131
7.2	Affine scale spaces . . . . .	136
7.2.1	Exact algorithm . . . . .	136
7.2.2	Simplified algorithm . . . . .	141
7.3	Affine scale space of non-closed curves . . . . .	157
<b>8</b>	<b>Conclusion</b>	<b>161</b>
8.1	Applications . . . . .	161
8.2	Further work . . . . .	162

<b>II</b>	<b>Multiscale Analysis of Movies for Depth Recovery</b>	<b>165</b>
<b>9</b>	<b>Introduction</b>	<b>167</b>
9.1	The depth recovery problem . . . . .	167
9.2	Geometric framework . . . . .	171
9.3	Velocity field . . . . .	172
9.4	Depth recovery . . . . .	173
<b>10</b>	<b>Axiomatic formulation</b>	<b>175</b>
10.1	Architectural axioms . . . . .	176
10.2	Specific axioms . . . . .	177
10.3	Fundamental equation . . . . .	178
<b>11</b>	<b>The Depth-Compatible Multiscale Analysis</b>	<b>185</b>
11.1	Classical solutions of the DCMA . . . . .	185
11.2	Weak solutions of the DCMA . . . . .	194
11.3	A viscosity formulation . . . . .	197
11.4	Appendix on the heat equation . . . . .	200
11.5	Further existence properties . . . . .	202
<b>12</b>	<b>Properties of the DCMA</b>	<b>203</b>
12.1	Checking the axioms . . . . .	203
12.2	Asymptotics of the DCMA . . . . .	205
12.3	Diffusion of the movement . . . . .	205
12.4	A conservation law . . . . .	210
12.4.1	Compactly supported movies . . . . .	210
12.4.2	Light Energy conservation . . . . .	211
12.5	A variational principle . . . . .	212
12.5.1	A minimization law . . . . .	212
12.5.2	A variational interpretation . . . . .	213
12.6	Interpretation for the observed scene . . . . .	214
12.6.1	Ideal movies . . . . .	214
12.6.2	Differential characterization of ideal movies . . . . .	215
12.6.3	Evolution of ideal movies . . . . .	216
12.6.4	Characterization of the DCMA . . . . .	218
<b>13</b>	<b>Numerical scheme and experiments</b>	<b>221</b>
13.1	Definition . . . . .	221
13.2	Consistency (regular case) . . . . .	222
13.3	Singular points . . . . .	228
13.4	Algorithms . . . . .	231
13.4.1	Data preparation . . . . .	231
13.4.2	Filtering with the DCMA . . . . .	231
13.4.3	Computing velocities . . . . .	232
13.5	Experiments . . . . .	234
13.5.1	TREES movie (natural) . . . . .	234
13.5.2	GARDEN movie (natural) . . . . .	240
13.5.3	Sensitivity to noise . . . . .	246

<b>14 Extensions and conclusion</b>	<b>251</b>
14.1 Extension to more general geometric configurations . . . . .	251
14.1.1 The camera motion is not horizontal . . . . .	251
14.1.2 The camera motion does not lie in the image plane . . . . .	251
14.1.3 Case of pure zooming . . . . .	252
14.2 Case of any rigid motion . . . . .	253
14.3 Occlusions . . . . .	254
14.4 Conclusion . . . . .	255
<b>Bibliography</b>	<b>257</b>

# Chapter 1

## Introduction

Le traitement d'images a fait son apparition dans les années 1970, lorsque le développement des ordinateurs a rendu possibles les premiers calculs numériques sur des images digitales. Pendant longtemps, il fut un domaine presque exclusivement réservé à des équipes d'ingénieurs motivés par des applications immédiates ; ce n'est que depuis quelques années qu'il a fait l'objet d'une étude mathématique plus rigoureuse, qui a permis de classifier bon nombre de techniques antérieures, et d'en expliquer le succès ou l'échec.

### 1.1 Problèmes et Enjeux du traitement d'images

Aujourd'hui encore, le traitement d'images (et de films) pose essentiellement trois grands problèmes : celui de **l'analyse** (comment obtenir des informations concernant les "objets" présents sur une image donnée ?), celui de **la restauration** (comment améliorer la qualité d'une image, la rendre plus nette ?) et celui de **la compression** (comment coder une image sous la forme la plus compacte possible ?). Nous ne parlerons pas ici de la synthèse d'images, qui ne relève pas à proprement parler du traitement d'images, même si elle lui est souvent complémentaire. L'analyse d'images est un maillon essentiel de la robotique, puisqu'elle doit définir les mécanismes de perception visuelle des robots. Elle a aussi beaucoup d'autres applications : reconnaissance de formes, cartographie aérienne, contrôle de qualité sur une chaîne de production, etc... La restauration d'images, quant à elle, est un outil très appréciable pour corriger des défauts qui apparaissent lors de la production d'une image : flou de focalisation ou flou de bougé, présence de "bruit", etc... Enfin, la compression d'images est un domaine relativement nouveau, devenu indispensable avec le développement massif des moyens de communication ; le débit de transmission d'un canal (hertzien ou filaire) étant toujours limité par des contraintes physiques, la compression apparaît comme un moyen simple d'augmenter le débit d'information. Malgré les apparences, ces trois problèmes ne sont pas indépendants : par exemple, le débruitage (suppression d'artefacts créés lors de l'acquisition) est une forme de restauration quasiment indispensable pour l'analyse. De même, le processus de simplification induit par l'analyse d'une image est une

étape préliminaire à certains algorithmes de compression. Enfin, certains algorithmes de compression avec perte d'information (cas de la norme JPEG par exemple) justifient une étape de restauration visant à corriger les défauts induits par la boucle compression-décompression. Dans toute cette étude, notre point de vue sera celui de l'analyse des images, même si nous aurons l'occasion d'illustrer quelques applications de ces procédés à la restauration.

## 1.2 Analyse des images : la notion de *scale space*

*Quelles informations peut-on extraire d'une image donnée ?* Cette question très générale est le point de départ de l'analyse d'images, qui précède souvent une phase d'interprétation. Par exemple, le fait que deux objets se déplacent à des vitesses différentes dans une séquence d'images est une information objective ; mais si l'on sait que ces deux objets sont en réalité fixes et que leur mouvement apparent est dû au déplacement de la caméra, on peut alors en déduire que l'objet qui se déplace le plus vite est plus proche de la caméra que l'autre, et même quantifier la distance de chaque objet au plan focal de la caméra.

**Historique.** Dans les années 1970, Bela Julesz mit en évidence l'existence de mécanismes réflexes dans la vision humaine. Il montra notamment que cette vision “bas niveau”, opérée dans les premières millisecondes du processus de vision, permet à l'homme de percevoir le relief et de discriminer des textures. Ses conclusions amenèrent les scientifiques à conjecturer l'existence d'un modèle mathématique simple de la vision préattentive. C'est ainsi que David Marr [53] formula quelques années plus tard le concept de “raw primal sketch”, ou pyramide visuelle. Selon lui, l'analyse visuelle résulte d'une représentation de l'image à des échelles différentes, allant d'une description très fine à l'échelle 0 vers une description de plus en plus globale et simplifiée à mesure que l'échelle augmente. Ainsi, une image brute est en réalité “vue” comme une collection d'images indexées par un paramètre d'échelle correspondant au degré de simplification opéré par rapport à l'image brute (échelle 0). Ce paramètre d'échelle peut d'ailleurs être identifié au temps d'analyse dans le système visuel humain. Cette représentation multiéchelle d'une image, qui fait donc intervenir une variable d'espace et une variable d'échelle, est appelée *scale space* [80] : son adéquation à modéliser la vision humaine préattentive a été bien vérifiée depuis, tant d'un point de vue psychophysique que biologique.

Le premier exemple de *scale space*, basé sur un filtrage linéaire, a rapidement montré ses limites. En effet, alors que la structure des appareils d'acquisition d'images rend, par la présence de filtres passe-bas, les théories linéaires bien adaptées à la compression (cf. la compression par ondelettes : [56], [28]), voire à la restauration (pour le déflouage par exemple), en revanche la nature “occlusive” des images à des échelles supérieures rend ces théories peu adaptées à l'analyse. En fait, le processus de formation d'une image naturelle<sup>1</sup> (un paysage, une scène

---

<sup>1</sup>c'est-à-dire issue d'un processus de vision, par opposition à une image scientifique utilisée pour représenter des données abstraites.

urbaine, ...) résulte d'un principe d'occlusion : lorsqu'un objet (non transparent) en cache un autre, seul le premier est visible, et l'on n'observe pas une espèce de superposition des images des deux objets. Cette constatation condamne immédiatement la généralisation hâtive des techniques linéaires employées en traitement du signal, pour l'analyse de la parole notamment, où il est clair que la nature même des ondes sonores implique un principe de superposition.

**Le principe morphologique.** Une alternative aux scale spaces linéaires apparut lors du développement de la Morphologie Mathématique ([70], [54]). Poussé à l'extrême, le principe morphologique s'énonce de la manière suivante : *“Dans une image, seul compte le fait qu'un point est plus clair ou plus foncé qu'un autre ; la valeur absolue de l'intensité n'est pas une information en soi.”* Cette hypothèse est d'ailleurs légitimée par l'exemple de la vision humaine : on ne voit pas différemment à travers une vitre teintée ! Concrètement, si l'on représente une image par une fonction  $u : \mathbb{R}^2 \rightarrow \mathbb{R}$  qui mesure en chaque point du plan l'intensité lumineuse reçue (le niveau de gris), ce principe dit que l'analyse de l'image  $u$  et d'une image de type  $g(u)$  (avec  $g : \mathbb{R} \rightarrow \mathbb{R}$  monotone) doit être la même. De tels changements de contraste  $g$ , qui opèrent une redistribution des niveaux de gris, sont de toute façon présents dans la chaîne d'acquisition. En pratique, le principe morphologique implique que le scale space opère indépendamment sur les lignes de niveau  $u = cte$  : l'analyse d'images se ramène donc à une analyse purement géométrique.

**Une classification axiomatique.** Peu à peu, de nouveaux modèles morphologiques sont apparus, et c'est en 1993 qu'une démarche axiomatique rigoureuse (cf. [4]) a permis de classifier complètement les théories existantes. Dans cette approche, chaque scale space est caractérisé en fonction de ses propriétés :

- Comment le scale space opère-t-il sur les niveaux de gris d'une image : est-il linéaire, morphologique, ... ?
- Avec quel groupe de transformations du plan le scale space commute-t-il : translations, rotations, symétries, affinités ?

D'autres propriétés, notamment le principe de comparaison (ou principe du maximum) qui assure qu'un scale space est un processus de simplification, avaient déjà été identifiées comme des propriétés fondamentales. De cette classification axiomatique, qui permit de regrouper au sein d'un même formalisme de nombreuses théories existantes, a alors émergé un nouveau modèle, baptisé *Affine Morphological Scale Space*. Ce scale space morphologique possède le groupe d'invariance le plus gros (en l'occurrence le groupe affine) que l'on puisse obtenir pour simplifier des images. La démarche axiomatique a aussi été appliquée en dimension supérieure, par exemple pour obtenir le premier scale space de films ([40]).



### 1.3 Plan de l'exposé.

Cette thèse est divisée en deux parties indépendantes. La première est consacrée à la reconnaissance de formes : nous élaborons et étudions un algorithme géométrique permettant de calculer avec précision le scale space affine d'une courbe plane. Dans la deuxième partie, nous rappelons comment, à partir d'un film, il est possible — en théorie — de retrouver le relief des objets apparaissant sur chaque image. Nous montrons ensuite comment le film doit être analysé pour qu'une telle opération soit effectivement réalisable.

**La reconnaissance de formes.** Le problème de la reconnaissance de formes peut être posé comme suit : *“Etant donné une base de formes de référence, comment reconnaître une nouvelle forme (ou la rejeter si elle n'est pas référencée dans la base), sachant qu'elle peut avoir été déformée et altérée ?”* Dans ce qui suit, nous entendrons par forme une région du plan délimitée par un nombre fini de courbes de Jordan. D'un point de vue pratique, l'approche morphologique permet d'extraire naturellement des formes d'une image en considérant simplement ses ensembles de niveau

$$\chi_\lambda(u) = \{\mathbf{x} \in \mathbb{R}^2, u(\mathbf{x}) \geq \lambda\}.$$

Grâce à cette décomposition, il devient alors équivalent de traiter une forme (c'est-à-dire un ensemble de courbes) et une image. Nous devons bien sûr préciser sous quelles conditions nous considérons deux formes comme semblables : par exemple, il semble naturel que la reconnaissance d'une forme ne dépende pas de sa position dans l'image. Mathématiquement, cela se traduit par l'identification d'un groupe de transformations du plan (contenant les translations d'après ce que nous venons de voir) qui induira une classe d'équivalence sur les formes.

Lorsque la reconnaissance est globale, un processus de normalisation peut être effectué : on choisit un représentant canonique dans la classe d'équivalence de chaque forme connue, et la reconnaissance se ramène alors à une comparaison entre deux représentants. Mais si l'on suppose que des masquages partiels peuvent intervenir, c'est à dire que les formes à reconnaître ne sont pas nécessairement “entières”, une telle normalisation devient impossible, à cause de la perte d'information induite par le masquage. Il faut alors tenter une reconnaissance locale, généralement basée sur un calcul de points caractéristiques (extremas de courbure, points d'inflexion, cf. [26]). Ces points caractéristiques, définis localement, impliquent souvent l'estimation de dérivées le long de la courbe, ce qui n'a en général pas de sens sur une courbe brute, dont le contour peut avoir été rendu très irrégulier par la présence de bruit dans l'image originale. Ainsi, un lissage préalable est nécessaire, et s'exprime naturellement sous la forme d'un scale space géométrique. Pour qu'un scale space géométrique soit effectivement un procédé de simplification, il faut qu'il satisfasse ce que l'on appelle le *principe d'inclusion locale* (cf. figure 1.1) : *“Si une forme est localement contenue dans une autre, alors cette inclusion doit persister localement pour une échelle d'analyse suffisamment petite.”* Ce principe est fondamental : il garantit notamment la

stabilité des algorithmes qui le vérifient. Formulé en termes d'images, il est alors équivalent au principe du maximum.

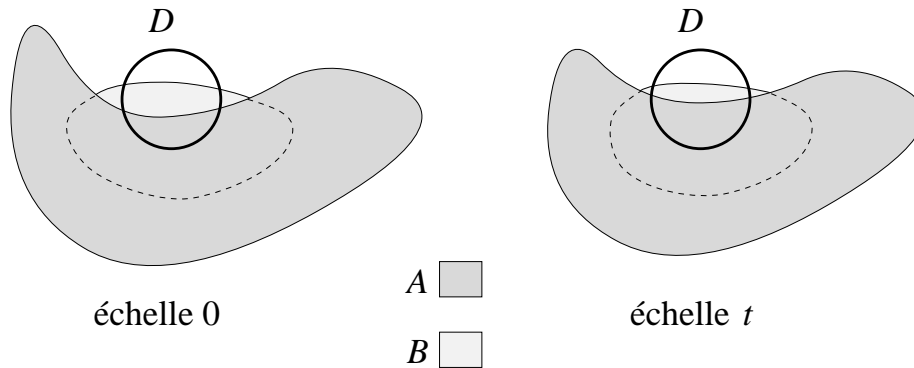


Figure 1.1: Illustration du principe d'inclusion locale.

*Une analyse multiéchelle vérifie le principe d'inclusion locale lorsque la condition suivante est vérifiée : “si une forme  $A$  est localement contenue dans une forme  $B$  (i.e  $A \cap D \subset B \cap D$  pour un certain voisinage  $D$ ), alors cette propriété reste vraie pour les formes analysées à une échelle suffisamment petite”. Ce principe garantit qu’une telle analyse multiéchelle est un processus de simplification, et assure la stabilité des algorithmes qui le vérifient.*

Moyennant une hypothèse supplémentaire de régularité relevant plus d’un artifice mathématique que d’une hypothèse physique, on obtient alors un unique modèle de scale space “maximal”, c’est à dire commutant avec le plus gros groupe possible de transformations du plan, en l’occurrence le groupe affine, engendré par les translations et les applications linéaires inversibles. Ce scale space affine, que nous avons évoqué précédemment, fut découvert simultanément deux équipes de chercheurs : L.Alvarez, F.Guichard, P.-L.Lions et J.-M.Morel [4] en termes d’analyse d’images, et par G.Sapiro et A.Tannenbaum dans sa version géométrique. Pour l’analyse de courbes, c’est cette dernière formulation qui est la plus adaptée : une courbe initiale  $s \mapsto \mathbf{C}(s, 0)$  évolue selon l’équation

$$\frac{\partial \mathbf{C}}{\partial t}(s, t) = \gamma(s, t)^{\frac{1}{3}} \mathbf{N}(s, t), \quad (1.1)$$

où  $\gamma(s, t)$  et  $\mathbf{N}(s, t)$  représentent respectivement la courbure et la normale au point  $\mathbf{C}(s, t)$  à la courbe  $s \mapsto \mathbf{C}(s, t)$ . La variable  $t$  représente le paramètre d’échelle que nous avons évoqué précédemment. Ainsi, la reconnaissance locale et affine-invariante de formes peut être effectuée de la manière suivante : on calcule le scale space affine à différentes échelles de la forme à reconnaître, puis on calcule des points caractéristiques sur la courbe obtenue à partir d’invariants affines locaux (la courbure affine par exemple) ou semi-locaux (cf. [26]). Enfin, on compare ces points caractéristiques avec ceux des formes de référence, ceci d’une manière affine invariante.

D’un point de vue numérique, la difficulté majeure qui intervient dans ce procédé est le calcul du scale space affine de la forme initiale. En effet, il n’existait jusqu’à présent qu’une manière raisonnable de calculer le scale space affine d’une forme, la méthode étant due à S.Osher et

J.A.Sethian. L'idée était de considérer l'image "fonction-distance" associée à la forme  $S \subset \mathbb{R}^2$  :

$$u(\mathbf{x}, 0) = \begin{cases} -\text{dist}(\mathbf{x}, \partial S) & \text{si } \mathbf{x} \in S, \\ +\text{dist}(\mathbf{x}, \partial S) & \text{si } \mathbf{x} \notin S, \end{cases}$$

et de lui appliquer un schéma aux différences finies pour calculer son scale space morphologique affine  $(u(\cdot, t))_{t \geq 0}$ . Grâce au principe morphologique, l'analyse de la forme  $S$  à l'échelle  $t$  était alors donnée par  $\{\mathbf{x} \in \mathbb{R}^2, u(\mathbf{x}, t) \leq 0\}$ .

Pourquoi le seul schéma "raisonnable" de scale space affine semblait-il nécessiter une formulation en termes d'images ? Simplement parce qu'un schéma géométrique aux différences finies basé sur une évolution de points est condamné à échouer : un tel schéma ne peut pas vérifier le principe d'inclusion, pourtant crucial pour assurer la stabilité de l'algorithme et sa consistance avec le scale space affine. En termes d'évolution d'images, le principe correspondant au principe d'inclusion (principe de comparaison, ou principe du maximum) est beaucoup plus facile à garantir numériquement, ce qui explique l'intérêt de la méthode d'Osher-Sethian. En revanche, l'invariance affine devient quelque peu illusoire, ne serait-ce qu'à cause de la grille fixe sur laquelle se placent les pixels de l'image. De plus, outre une certaine lourdeur, ce procédé est limité dans la précision de ses résultats à cause du pas de la grille sous-jacente.

Dans la première partie de cette thèse, nous proposons une alternative géométrique à la méthode d'Osher-Sethian pour calculer numériquement le scale space affine d'une courbe. Le schéma que nous décrivons est basé sur l'itération d'un opérateur **géométrique** et **non-local** (cf. figure 1.2), qui vérifie le principe d'inclusion et l'invariance affine, et qui peut être calculé facilement pour des courbes polygonales. L'algorithme obtenu est rapide, stable et très précis, comme l'illustrent les nombreuses expériences que nous effectuons.

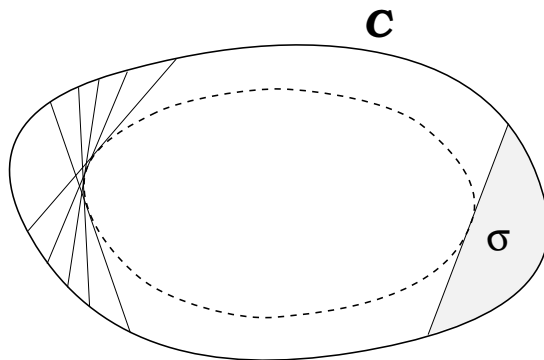


Figure 1.2: "Erosion affine" d'une courbe convexe.

*Représentée en pointillés, l'érosion affine de paramètre  $\sigma$  de la courbe convexe  $\mathcal{C}$  est obtenue en éliminant de l'intérieur de  $\mathcal{C}$  toutes les régions d'aire  $\sigma$  délimitées par un arc de courbe et une corde de  $\mathcal{C}$ . Cela revient en général à considérer l'enveloppe (ou encore les milieux) de telles cordes. En itérant un tel opérateur, on obtient alors une approximation fine et numériquement stable du scale space affine de  $\mathcal{C}$ .*

**La détermination du relief.** Comme nous l'avons évoqué auparavant, l'un des enjeux majeurs de la robotique est de résoudre le problème de la perception. Lorsqu'un robot se déplace dans un environnement connu ou inconnu, il lui est nécessaire de pouvoir se repérer, voire d'établir une carte tridimensionnelle du monde qui l'entoure. Ce genre d'opération est effectué en permanence par le système visuel humain, qui utilise conjointement différents types d'informations. Par exemple, la quantité de lumière réfléchiée en chaque point d'une surface donne en général une information sur la direction de la normale à cette surface : c'est le *shape from shading*. L'utilisation de la stéréovision, basée sur l'analyse des petites différences entre les images reçues par chaque oeil, permet aussi de déterminer le relief. Mais même lorsque l'on ferme un oeil, notre système visuel reconstruit sans problème le relief observé pour peu que l'on se déplace un peu : le mouvement apparent des objets nous renseigne sur leur distance effective, les objets les plus près étant animés d'un mouvement apparent plus rapide, alors que les objets très éloignés restent quasiment fixes. D'autres informations, relevant de connaissances a priori, sont aussi couramment utilisées par le système visuel humain : connaissant la taille approximative d'une voiture, nous pouvons facilement déduire de sa taille apparente la distance à laquelle elle se trouve. Ce type de perception du relief est cependant beaucoup plus complexe, et ne résulte pas d'un processus de vision préattentive, contrairement à la stéréovision par exemple.

Les premiers essais de reconstruction automatique du relief furent basés en toute logique sur le procédé de stéréovision. A partir de deux images obtenues grâce à deux caméras légèrement décalées, il semblait possible de reconstruire entièrement le relief de la scène observée. Bien que correcte d'un point de vue théorique, cette méthode se heurta assez rapidement à deux problèmes majeurs. Le premier, structurel, fut mis en évidence par un calcul simple montrant l'impossibilité d'obtenir à la fois un algorithme robuste (la comparaison des deux images est d'autant plus facile que les deux caméras sont proches) et une bonne estimation du relief (cet estimation est d'autant plus précise que les deux caméras sont éloignées). Le deuxième problème majeur survint à cause des techniques développées pour comparer les deux images : généralement basée sur l'extraction de contours rectilignes fortement contrastés, la comparaison n'est vraiment efficace que pour des scènes artificielles (bâtiments, routes, machines, ...), et ses performances chutent complètement dans le cas de scènes naturelles où des textures apparaissent plutôt que des arêtes vives (champs, herbe, feuillage, ...).

Ainsi, il était naturel de se tourner vers un procédé plus robuste, le *shape from motion*. L'idée est de considérer non plus deux images mais une séquence de plusieurs dizaines d'images, et d'analyser le déplacement apparent des objets entre chaque couple d'images successives, sous l'hypothèse que les objets sont en réalité fixes et que leur mouvement apparent n'est dû qu'au déplacement de la caméra. Une telle approche est en quelque sorte une généralisation du principe de stéréovision, et l'on devine que la redondance de l'information disponible (théoriquement, chaque couple d'images de la séquence produit une interprétation du relief d'après le principe de stéréovision) doit permettre de vaincre les problèmes de robustesse et de précision inhérents

au procédé de stéréovision. Si une telle approche semble en effet beaucoup plus fiable, de nouveaux problèmes apparaissent néanmoins. Dans l'approche continue généralement adoptée pour représenter la séquence d'images obtenue, la reconstruction du relief implique le calcul d'un rapport de dérivées, qui s'avère être très instable à cause de l'irrégularité spatiale des images et de leur trop rapide évolution temporelle tout au long de la séquence. D'autre part, la redondance de l'information contenue dans la séquence ne peut être pleinement exploitée que par une analyse globale de la séquence, et non par le calcul de dérivées à un instant donné.

Dans la deuxième partie de cette thèse, nous montrons comment, sous l'hypothèse que le mouvement de la caméra est donné, le problème de l'analyse du relief peut être résolu à l'aide d'un filtrage adéquat de la séquence d'images, qui permet d'induire une cohérence temporelle globale entre toutes les images de la séquence, ramenant ainsi l'analyse globale du relief à un calcul local. Ce processus de filtrage — un scale space faisant intervenir le temps, l'espace et l'échelle — est même unique, caractérisé par un ensemble de propriétés imposées par la géométrie du problème de la détermination du relief. Si le mouvement est donné par la variable spatiale  $x$ , l'analyse d'un film  $u(x, y, \theta, 0)$  ( $\theta$  représentant le temps et la dernière coordonnée l'échelle  $t$ ) est décrit par l'équation d'évolution<sup>2</sup>

$$u_t = u_{\theta\theta} - 2 \frac{u_\theta}{u_x} u_{\theta x} + \left( \frac{u_\theta}{u_x} \right)^2 u_{xx}. \quad (1.2)$$

Cette équation aux dérivées partielles non linéaire du second ordre, parabolique dégénérée, présente une singularité très forte lorsque la dérivée  $u_x$  s'annule, ce qui l'empêche de relever de la théorie classique des solutions de viscosité (cf. [27]), seule théorie de solutions faibles a priori adaptée à ce type d'équation. Néanmoins, nous établissons des résultats d'existence et d'unicité pour (1.2), et mettons en évidence certaines de ses propriétés qui se prêtent facilement à une interprétation physique. En particulier, il apparaît que ce scale space est vraiment compatible avec la reconstruction du relief, puisqu'il préserve tout film idéal, c'est-à-dire possédant déjà une interprétation cohérente en termes de relief observé et de mouvement de la caméra. Nous décrivons ensuite un schéma numérique pour résoudre (1.2), basé sur l'itération d'opérateurs morphologiques de type inf-sup. Enfin, par quelques expériences, nous confirmons numériquement les effets de cet équation : l'établissement d'une cohérence globale entre toutes les images du film qui ramène le calcul du relief à un processus simple et fiable.

---

<sup>2</sup>Selon la convention habituelle, les indices désignent des dérivées partielles.

# Bibliography

- [1] J.Aloimonos, A.Bandyopadhyay, “Active Vision”, Proceedings IEEE 1987, pp. 35-54.
- [2] S.J.Altschuler, M.A.Grayson, “Shortening space curves and flow through singularities”, *Journal of Differential Geometry* 35, pp. 283-298, 1992.
- [3] L.Alvarez, “Axiomatisation et Nouveaux Opérateurs de la Morphologie Mathématique”, Preprint CEREMADE, 1992.
- [4] L. Alvarez, F. Guichard, P.L. Lions, J.M. Morel, “Axioms and fundamental equations of image processing”, *Archives for Rational Mechanics* 123, pp. 199-257, 1993.
- [5] L. Alvarez, F. Morales, “Affine Morphological Multiscale Analysis of Corners and Multiple Junctions”, *International Journal of Computer Vision*, 1995.
- [6] L. Alvarez, J.-M. Morel, “Formalization and Computational Aspects of Image Analysis”, *Acta Numerica*, Cambridge University Press, 1994.
- [7] V.Arnold, *Ordinary Differential Equations*, Springer Verlag, 1992.
- [8] M.Asada, “Cylindrical Shape from Contour and Shading without knowledge of lighting conditions or surface albedo”, Proceedings IEEE 1987, pp. 412-416.
- [9] G.Barles, C.Georgelin, “A simple proof for the convergence of an approximate scheme for computing Mean Curvature Motions”, *SIAM Journal of Numerical Analysis* 32:2, pp. 484-500, 1995.
- [10] G.Barles, *Solutions de viscosité des équations de Hamilton-Jacobi*, Springer Verlag, Berlin, 1994.
- [11] J.L.Barron, D.J.Fleet, S.S.Beauchemin “Performance of Optical Flow Techniques”, *International Journal of Computer Vision* 12:1, pp. 43-77, 1994.
- [12] J.Bence, B.Merriman, S.Osher, “Diffusion generated motion by mean curvature”, Preprint, 1992.

- [13] R.C.Bolles, H.H.Baker, D.H.Marimont, "Epipolar-Plane Image Analysis : An Approach to Determining Structure from Motion", *International Journal of Computer Vision* 1, pp. 7-55, 1987.
- [14] S.Boukir, F.Chaumette "Reconstruction 3D de primitives géométriques par vision active", INRIA report #1722, july 1992.
- [15] M.Brady, A.Yuille "An Extremum Principle for Shape from Contours", Proceedings IEEE, pp. 186-199, 1984.
- [16] Arne Broman, *Introduction to Partial Differential Equations*, Dover, New-York, 1989.
- [17] A.M. Bruckstein, A. Cohen, G. Sapiro, "A Subdivision Scheme for continuous-scale B-splines and affine invariant progressive smoothing", Preprint CEREMADE #9403, 1994.
- [18] A.M.Bruckstein, A.Tannenbaum, "Some Mathematical Problems in Computer Vision", Preprint, Israel Institute of Technology, july 1992.
- [19] P.Burlina, R.Chellappa, "Spatio-temporal moments and generalized spectral analysis of divergent images for motion estimation". Proceedings IEEE, pp. 328-332, 1994.
- [20] F.Cao, "Morphological and Affine Invariant Filters : the N-dimensionnal case of Infsup Operators", to appear.
- [21] V.Caselles, B.Coll, J.-M. Morel "Is Scale Space possible ?", to appear in *IEEE Transactions on Image Processing*.
- [22] F. Catté, F. Dibos, "A morphological approach of mean curvature motion", Preprint CEREMADE #9310, 1993.
- [23] Y.G.Chen, Y.Giga, S.Goto, "Uniqueness and existence of viscosity solutions of generalized Mean Curvature flow Equations", *Journal of Differential Geometry* 33, pp. 749-786, 1991.
- [24] D.L.Chopp, "Numerical Computation of Self-Similar Solutions for Mean Curvature Flow", Internal Report, 1993.
- [25] R.Cipolla, A.Blake, "Surface Shape from the Deformation of Apparent Contours", *International Journal of Computer Vision* 9:2, pp. 83-112, 1992.
- [26] T.Cohignac, "Reconnaissance de formes planes", PhD Dissertation, CEREMADE, 1994.
- [27] M.G.Crandall, H.Ishii,P.-L. Lions, "User's guide to viscosity solutions of second order partial differential equations", *Bulletin of American Mathematical Society* 27, pp. 1-67, 1992.
- [28] J.-P.d'Alès, "Approximation linéaire et non linéaire de fonctions aléatoires : application à la compression d'images", PhD Dissertation, CEREMADE, 1996.

- [29] C.L.Epstein, M.Gage, “The curve shortening flow”, *Wave Motion Theory, Modeling and Computation*, A.Chorin and A.Majda Editors, Springer-Verlag, New-York, 1987.
- [30] L.C.Evans, “Convergence of an Algorithm for Mean Curvature Motion”, Preprint, 1992.
- [31] L.C.Evans, J.Spruck, “Motion of Level Sets by Mean Curvature I”, *Journal of Differential Geometry* 33, pp. 635-681, 1991.
- [32] L.C.Evans, “A geometric Interpretation of the Heat Equation with Multivalued Initial Data”, *SIAM Journal of Mathematical Analysis* 27:2, pp. 932-958, 1996.
- [33] F.Falzon, G.Giraudon, M.Berthod, “Vers un modèle pour l’analyse multiéchelle de surfaces 3D”, INRIA report #1639, march 1992.
- [34] O.Faugeras, T.Papadopoulos, “A theory of the motion fields of curves”, *International Journal of Computer Vision* 10:2, pp. 125-156, 1993.
- [35] O.Faugeras, Lustman, Toscani, “Motion and Structure from Motion from Point and Lines Matches”, Proceedings IEEE, pp. 25-34, 1987.
- [36] M. Gage and R.S. Hamilton, “The heat equation shrinking convex plane curves”, *Journal of Differential Geometry* 23, pp. 69-96, 1986.
- [37] Y.Giga, S.Goto, “Motion of hypersurfaces and geometric equations”, *Journal of Mathematical Society Japan*, vol. 44, no 1, 1992.
- [38] E. Giusti, *Minimal surfaces and functions of bounded variation*, Birkhäuser, 1984.
- [39] M.A. Grayson, “The heat equation shrinks embedded plane curves to round points”, *Journal of Differential Geometry* 26, pp. 285-314, 1987.
- [40] F. Guichard, “Axiomatisation des analyses multi-échelles d’images et de films”, PhD Dissertation, CEREMADE, 1994.
- [41] F. Guichard and J.-M. Morel, “Partial differential equations and image iterative filtering”, Preprint CEREMADE and Tutorial ThP2 of *IEEE International Conference on Image Processing*, 1995.
- [42] D.J.Heeger, “Optical Flow using Spatiotemporal filters”, *International Journal of Computer Vision*, pp. 279-302, 1988.
- [43] B.Horn, B.Shunk, “Determining Optical Flow”, *Artificial Intelligence* 17, pp. 185-203, 1981.
- [44] B.Horn, “Motion Fields Are Hardly Ever Ambiguous”, *International Journal of Computer Vision* 1, pp. 259-274, 1987.



- [45] H.Ishii, "A generalization of the Bence, Merriman and Osher algorithm for motion by mean curvature", Preprint.
- [46] T.Kanade, "Geometrical Aspects of Interpreting Images as a 3D scene", Proceedings IEEE 1983, pp. 555-568.
- [47] G.Kanizsa, *Grammatica del Vedere*, Il Mulino, Bologna, 1980. Traduction : *La grammaire du voir*, Diderot Editeur, Arts et Sciences, 1996.
- [48] R.Kimmel, A.M.Bruckstein, "Shape Offsets via Level sets", Preprint, Israel Inst. of Technology, march 1992.
- [49] G.Koepfler, C.Lopez, J.-M.Morel, "A Multiscale Algorithm for image segmentation by variational method", *Siam Journal of Numerical Analysis* 31:1, pp. 282-299, 1994.
- [50] H.Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections", *Nature* vol. 293, no. 5828, pp. 133-135, 1981.
- [51] C.Lopez, J.-M.Morel, "Axiomatisation of Shape Analysis and Application to Texture Hyperdiscrimination". Preprint CEREMADE, dec. 1992.
- [52] L.Marce, P.Bouthemy, "Determination of a depth map from an image sequence", Rapport INRIA #765, dec. 1987.
- [53] D.Marr, *Vision*, Freeman and Co., 1982.
- [54] G.Matheron, *Random Sets and Integral Geometry*, John Wiley, New York, 1975.
- [55] F.Meyer, P.Bouthemy, "Region-Based Tracking in an Image Sequence", INRIA report #1723, july 1992.
- [56] Y.Meyer, *Les ondelettes, algorithmes et applications*, Armand Collin, Paris, 1992.
- [57] L.Matthies, T.Kanade, R.Szeliski, "Kalman Filter-based Algorithms for Estimating Depth from Image Sequences", *International Journal of Computer Vision* 3, pp. 209-236, 1989.
- [58] L.Moisan, "Analyse multiéchelle de films pour la reconstruction du relief", Note au compte rendu de l'Académie des sciences, Paris, tome 320, série I, pp. 279-284, feb. 1995.
- [59] L.Moisan, "Perspective invariant movie analysis for depth recovery", Proceedings SPIE, vol. 2567, pp. 84-94, july 1995.
- [60] L.Moisan, "Multiscale Analysis of Movies for Depth Recovery", *Proceeding of the International Conference on Image Processing*, vol. 3, pp. 25-28, oct. 1995.
- [61] L.Moisan, "Affine plane curve evolution : a fully consistent scheme", Preprint CEREMADE, nov. 1996, to appear in *IEEE Transactions on Image Processing*.

- [62] P.Monasse, “A review of Optical Flow Techniques”, Personal communication, 1997.
- [63] D.Mumford, “The problem of Robust Shape Descriptors”, *Proceedings IEEE*, pp. 602-606, 1987.
- [64] A.Naeve, J.-O.Eklundh, “On Projective Geometry and the Recovery of 3D Structure”, *Proceedings IEEE*, pp. 128-135, 1987.
- [65] S.Osher and J.A.Sethian, “Fronts Propagating with Curvature-Dependent Speed : Algorithms Based on Hamilton-Jacobi Formulations”, *Journal of Computational Physics* 79:1, pp. 12-49, 1988.
- [66] A.Rosenfeld, “Image Analysis : Problems, progress and prospects”, *Pattern Recognition* 17:1, pp. 3-12, 1984.
- [67] G.Sapiro and A.Tannenbaum, “On affine plane curve evolution”, *Journal of Functional Analysis* 119, pp. 79-120, 1994.
- [68] G.Sapiro and A.Tannenbaum, “Affine Invariant Scale Space”, *International Journal of Computer Vision* 11:1, pp. 25-44, 1993.
- [69] M.Schmitt, *Morphologie Mathématique*, Masson, 1993.
- [70] J.Serra, *Image analysis and mathematical morphology*, vol. 2, Academic Press, 1982.
- [71] J.A. Sethian, “Curvature and the Evolution of Fronts”, *Communications in Mathematical Physics* 101, pp. 487-499, Springer-Verlag, 1985.
- [72] J.A. Sethian, *Level Set Methods*, Cambridge University Press, 1996.
- [73] S.Soatto, R.Frezza, “Recursive Motion Estimation on the Essential Manifold”, Rapport interne, 1993.
- [74] M.Spetsakis, J.Aloimonos, “A multi-frame Approach to Visual Motion Perception”, *International Journal of Computer Vision* 6:3, pp. 245-255, 1991.
- [75] K.A.Stevens, “The Visual Interpretation of Surfaces Contours”, *Artificial Intelligence* 17, pp. 47-73, 1981.
- [76] G.Tziritas, “Estimation of motion and structure of 3D objects from a sequence of images”, *Proceedings IEEE*, pp. 693-697, 1987.
- [77] E.J.Weldon Jr., H.Liu, “How accurately can direct motion vision determine depth ?”, *Proceedings IEEE*, pp. 613-618, 1991.
- [78] J.Weng, “Optimal Motion and Structure Estimation”, *Pattern Analysis and Machine Intelligence* 15:9, pp. 864, sept. 1993.

- [79] M.Wertheimer, "Untersuchungen zur Lehre der Gestalt", II, Psychologische Forschung, vol. 4, pp. 301-350, 1923.
- [80] A.P.Witkin, "Scale-Space filtering", Proceedings of IJCAI, Karlsruhe, pp. 1019-1021, 1983.
- [81] G.Xu, S.Tsuji, "Inferring Surfaces from boundaries", Proceedings IEEE, pp. 716-720, 1987.