

ACADEMIE DE MONTPELLIER  
**UNIVERSITE MONTPELLIER II**  
SCIENCES ET TECHNIQUES DU LANGUEDOC

**T H E S E**

présentée à l'Université de Montpellier II Sciences et Techniques du Languedoc  
pour obtenir le diplôme de DOCTORAT

**Spécialité** : INFORMATIQUE

**Formation Doctorale** : Informatique

**Ecole Doctorale** : Sciences pour l'Ingénieur

**ANALYSE ET REPRÉSENTATION  
DE SCÈNES COMPLEXES  
PAR GROUPEMENT PERCEPTUEL**

Application à la perception de structures curvilignes.

par

Laurent ALQUIER

Soutenue le 30 septembre 1998 devant le Jury composé de :

M. HABIB Michel	Professeur, LIRMM Montpellier	Président
Mme ZERUBIA Josiane	Directeur de Recherche, INRIA Sophia Antipolis	Rapporteur
M. JOLION Jean-Michel	Professeur, INSA Lyon	Rapporteur
M. FIORIO Christophe	Maître de conférence, LIRMM Montpellier	Examineur
Mme. PHILIPP Sylvie	Professeur, ENSEA Paris	Examineur
M. OUSSALAH Chabane	Professeur, LGI2P Nîmes	Directeur de Thèse
M. MONTESINOS Philippe	Maître assistant, LGI2P Nîmes	Responsable de Thèse

A Lori, pour son infinie patience.  
A Linfa, pour avoir quitté l'une de ses neuf vies si tôt.



Je tiens à exprimer ma reconnaissance aux personnes qui ont contribué, de près ou de loin, à l'accomplissement de ce travail. En particulier, je voudrais remercier tout d'abord ceux qui m'ont fait l'honneur de participer au jury de cette thèse.

Monsieur Mourad Oussalah, Professeur à l'École des Mines d'Alès, pour la confiance qu'il m'a accordée en acceptant de diriger ce travail, et Madame Janine Magnier, Directeur du Laboratoire de Génie Industriel et d'Ingénierie de Production pour m'avoir accueilli dans son laboratoire.

Monsieur Philippe Montesinos, Maître assistant au LGI2P, pour son accueil au sein de l'équipe Vision du laboratoire. Il a su par ses conseils, sa disponibilité et les fructueuses discussions que nous avons eues ensemble, me communiquer sa passion pour la vision par ordinateur.

Madame Josiane Zerubia, Directeur de Recherche à l'INRIA et Monsieur Jean-Michel Jolion, Professeur à l'INSA de Lyon, pour le temps accordé à juger ce travail et les conseils précieux qu'ils m'ont donnés pour l'améliorer.

Monsieur Michel Habib, Professeur à l'Université Montpellier II, Monsieur Christophe Fiorio, Maître de Conférence au LIRMM, et Madame Sylvie Philipp, Professeur à l'ENSEA de Paris, pour l'intérêt qu'ils ont porté à mon travail en acceptant de faire partie du jury.

Toute ma sympathie pour les membres du laboratoire, enseignants-chercheurs et doctorants, ainsi que l'ensemble du personnel du site EERIE pour la qualité de leur accueil. Merci, notamment, à Sylvie, Gérard, Mireille, Michel, Vincent, et les autres...

Les membres de l'équipe Vision, les nouveaux comme les anciens, pour tout ce que j'ai pu apprendre à leur contact. En particulier, Nasser Armande pour ses contributions indirectes à ce travail, Houda Chabbi pour ses précieux conseils de rédaction et Pierre Soille pour de nombreuses discussions stimulantes.

Mes deux "cyber-mamans" : Françoise, pour ses encouragements en période de stress, pour avoir subi la relecture des premières versions de ce mémoire et surtout, pour les mercredi matin de la bib'. Annie pour sa confiance, sa bonne humeur contagieuse et son soutien pendant les longs séjours au Centre de Calcul. Vive le F.A.A.L!

Les piliers du Centre de Calcul : François, pour son aide technique inestimable, ses conseils de programmation et sa disponibilité, devant une machine comme autour d'une bonne table. Nicolas et Fabien pour les longues soirées de travail à la lueur des néons.

Sylvain et Christelle pour leur profonde amitié, les nombreuses pauses café et les longues discussions à refaire le monde (en particulier celui des D'ni). Sans oublier Zoé, voyageuse nocturne, grâce à qui j'ai pu rédiger cette thèse sur mon propre ordinateur.

Mes parents et mes proches pour leurs encouragements, en particulier Fabien pour avoir pris le temps de pousser son grand frère à aller voir ailleurs que ses bouquins et l'écran de sa machine.

Enfin Lori, pour m'avoir attendu depuis si longtemps.



---

# Sommaire

---

<b>Introduction</b>	<b>1</b>
<b>I Perception de structures curvilignes en vision par ordinateur</b>	<b>5</b>
<b>1 Vision naturelle et artificielle</b>	<b>7</b>
1.1 Problématique de la vision par ordinateur . . . . .	7
1.2 Théories de la perception visuelle . . . . .	9
1.2.1 Comment définir la vision? . . . . .	9
1.2.2 Psychologie de la perception visuelle . . . . .	11
1.2.3 Théories neuro-physiologistes . . . . .	20
1.3 Méthodologies de la vision par ordinateur . . . . .	22
1.3.1 Paradigme reconstitutif . . . . .	22
1.3.2 Approche “intentionnelle” . . . . .	28
1.3.3 Approche “système” . . . . .	32
1.4 Conclusion . . . . .	37
<b>2 Analyse et interprétation des scènes de contours</b>	<b>39</b>
2.1 Représentation de scènes complexes . . . . .	39
2.1.1 Acquisition des images . . . . .	41
2.1.2 Segmentation d’indices visuels . . . . .	41
2.1.3 Structuration des indices visuels . . . . .	42
2.1.4 Représentation de haut niveau . . . . .	45
2.2 Scènes de contours . . . . .	47
2.2.1 Scènes de contours et représentations . . . . .	48
2.2.2 Sources d’ambiguïtés et contraintes . . . . .	50
2.3 Définition et détection des contours . . . . .	52
2.3.1 Discontinuité d’intensité . . . . .	52
2.3.2 Modèles de contours actifs . . . . .	57
2.3.3 Coins, sommets . . . . .	59
2.3.4 Réseaux fins . . . . .	59
2.3.5 Frontières de régions homogènes . . . . .	61

2.3.6	Contours fictifs . . . . .	62
2.4	Structuration des contours . . . . .	63
2.4.1	Structuration directe . . . . .	64
2.4.2	Structuration progressive . . . . .	65
2.5	Hauts niveaux de représentations . . . . .	67
2.5.1	Représentations bi-dimensionnelles . . . . .	67
2.5.2	Représentations tri-dimensionnelles . . . . .	68
2.5.3	Représentations sans reconstruction . . . . .	72
2.6	Conclusion . . . . .	74
<b>3</b>	<b>Groupement perceptuel en vision par ordinateur</b>	<b>75</b>
3.1	Principes d'organisation perceptuelle . . . . .	75
3.2	Règles de groupement . . . . .	78
3.3	Application à la vision par ordinateur . . . . .	84
3.3.1	Principe de "non-accidentalité" . . . . .	85
3.3.2	Choix des règles de groupements . . . . .	86
3.3.3	Choix des primitives à grouper . . . . .	87
3.4	Techniques de groupement de contours . . . . .	88
3.4.1	Approches algorithmiques . . . . .	89
3.4.2	Méthodes d'optimisation . . . . .	90
3.4.3	Théorie des graphes . . . . .	92
3.4.4	Autres approches . . . . .	93
3.5	Principes de notre méthode . . . . .	94

## **II Analyse de scènes de contours par groupement perceptuel** **99**

<b>4</b>	<b>Saillance structurelle et groupements élémentaires</b>	<b>101</b>
4.1	Saillance structurelle . . . . .	101
4.2	Méthodologie de groupement par réseaux de saillance . . . . .	103
4.2.1	Définitions . . . . .	104
4.2.2	Voisinage . . . . .	105
4.2.3	Optimisation par programmation dynamique . . . . .	108
4.2.4	Groupement à partir d'une carte de saillance . . . . .	114
4.2.5	Conclusion et applications . . . . .	120
4.3	Application au groupement de pixels . . . . .	122
4.3.1	Primitive "pixel" . . . . .	122
4.3.2	Voisinage statique . . . . .	122
4.3.3	Fonction de qualité . . . . .	124
4.3.4	Mesure de saillance . . . . .	127
4.3.5	Optimisation . . . . .	128
4.3.6	Extraction et sélection des meilleurs groupes . . . . .	129

4.3.7	Résultats et développements éventuels . . . . .	131
4.4	Application au groupement de chaînes . . . . .	141
4.4.1	Primitive “chaîne” . . . . .	141
4.4.2	Voisinage dynamique . . . . .	142
4.4.3	Fonction de qualité . . . . .	148
4.4.4	Mesure de saillance . . . . .	150
4.4.5	Optimisation et sélection des meilleures courbes . . . . .	151
4.4.6	Résultats et perspectives . . . . .	152
<b>5</b>	<b>Eléments de représentation et groupements intermédiaires</b>	<b>161</b>
5.1	Structuration hiérarchique . . . . .	161
5.1.1	Analyse des groupements élémentaires . . . . .	162
5.1.2	Principes de groupement intermédiaire . . . . .	162
5.2	Hypothèses “segments” . . . . .	165
5.2.1	Détection des segments . . . . .	165
5.2.2	Organisation perceptuelle des segments . . . . .	169
5.3	Hypothèses “arcs” . . . . .	185
5.3.1	Détection d’arcs élémentaires . . . . .	185
5.3.2	Organisation perceptuelle des arcs . . . . .	193
5.4	Conclusions et perspectives . . . . .	200
<b>6</b>	<b>Groupements de haut niveau et mise en correspondance</b>	<b>203</b>
6.1	Mise en correspondance structurelle . . . . .	203
6.1.1	Relations structurelles . . . . .	205
6.1.2	Organisation perceptuelle et mise en correspondance . . . . .	206
6.2	Extraction et groupement de jonctions . . . . .	208
6.2.1	Détection des jonctions élémentaires . . . . .	209
6.2.2	Groupement en jonctions complexes . . . . .	210
6.2.3	Résultats sur les jonctions de segments . . . . .	212
6.3	Mise en correspondance de jonctions . . . . .	219
6.3.1	Coopération entre appariement et groupement . . . . .	220
6.3.2	Mesures de distances entre jonctions . . . . .	223
6.3.3	Relaxation “temporelle” . . . . .	229
6.3.4	Relaxation “perceptuelle” . . . . .	231
6.3.5	Résultats de mise en correspondance . . . . .	233
6.4	Conclusion . . . . .	239
<b>7</b>	<b>Conclusion</b>	<b>241</b>
	<b>Annexes</b>	<b>245</b>



<b>A Réseaux de saillance de Shashua et Ullman</b>	<b>247</b>
A.1 Mesures de saillance structurelle . . . . .	247
A.1.1 Optimisation combinatoire . . . . .	247
A.1.2 Mesures 'directes' . . . . .	251
A.2 Réseaux de Saillance de Shashua et Ullman . . . . .	254
A.2.1 Définitions et notations . . . . .	255
A.2.2 Mesure de saillance . . . . .	256
A.2.3 Optimisation récursive . . . . .	260
A.2.4 Extraction des structures saillantes . . . . .	260
A.2.5 Discussion . . . . .	262
<b>B Résultats complémentaires</b>	<b>267</b>
<b>Bibliographie</b>	<b>293</b>
<b>Index</b>	<b>309</b>

# Introduction

---

*“On aura compris qu’il n’y a pas d’un côté l’Image, matériau unique, inerte et stable, et de l’autre le Regard, comme un rayon de soleil mobile qui viendrait animer la page d’un livre grand ouvert. Regarder n’est pas recevoir mais ordonner le visible, organiser l’expérience. L’image tire son sens du regard, comme l’écrit de la lecture, et ce sens n’est pas spéculatif mais pratique.”*

“Vie et mort de l’image” - Régis Debray [Debray, 1992]



**Figure 0.1** - *Marché d'esclaves avec buste invisible de Voltaire* - Salvador Dalí  
Marché d'esclaves avec buste invisible de Voltaire - 1940 - Salvador Dalí Museum,  
San Petersburg, Floride

La perception des objets d'une scène, en particulier leur forme, est l'un des problèmes fondamentaux de la vision par ordinateur. L'acquisition d'images monoculaires reste la méthode la plus économique pour percevoir une scène. Mais elle est également source de nombreux problèmes mal posés, liés aux ambiguïtés créées par la projection de la scène sur l'image. En d'autres termes, une image seule ne contient pas assez d'informations pour une interprétation complète et sans ambiguïté de la scène représentée.

Pourtant, la vision humaine démontre l'existence de mécanismes robustes pour la perception de structures en trois dimensions, même lorsque la scène est perçue à partir d'une simple image. Le système visuel humain présente en particulier de nombreux mécanismes dont le rôle est de guider en permanence la perception dans un flot continu d'informations visuelles. Un certain nombre de ces mécanismes ont été appliqués à la vision par ordinateur. Inspirés de la vision naturelle, l'utilisation de vues multiples, de focus d'attention ou encore de rétine artificielle à résolution polaire ont montré leur intérêt pour résoudre certaines ambiguïtés de la recherche visuelle. D'autre part, des théories issues de la psychologie de la perception, fournissent des contributions importantes. C'est le cas, en particulier, du groupement perceptuel.

Le groupement perceptuel désigne la faculté de la vision humaine à organiser certains éléments visuels en groupes significatifs. Ce phénomène se produit de façon spontanée, avant toute interprétation du contenu de la scène observée. Des expériences psycho-visuelles simples montrent clairement l'importance de certaines relations, telles que la proximité, la fermeture, la continuité ou la symétrie, lors de ces groupements.

L'organisation perceptuelle a, depuis une vingtaine d'années, fait l'objet de nombreuses recherches en vision par ordinateur. Ce problème combinatoire complexe est le plus souvent exprimé comme le groupement d'attributs détectés dans l'image selon des structures régulières plus complexes. Contrairement aux méthodes classiques d'analyse d'images, ce type d'approche met l'accent sur des critères qualitatifs et génériques, au lieu de rechercher des modèles mathématiques précis. Finalement, ces groupements sont utilisés pour initialiser plus facilement des traitements de haut niveau tels que la reconnaissance de formes ou bien la mise en correspondance d'images.

Au cours de notre travail, nous nous sommes intéressés au rôle que peut jouer le groupement perceptuel dans la perception des contours des objets d'une scène. Plus précisément, l'objectif de cette thèse est d'extraire, à partir d'une détection de contours, les éléments caractéristiques des principales structures curvilignes de la scène. Il s'agit donc, dans un premier temps, de définir des critères de qualité permettant de mettre en valeur les contours les plus saillants de l'image. Ces contours dominants servent de point de départ à l'extraction des éléments caractéristiques de la scène.

Nous proposons une approche hiérarchique pour extraire les éléments visuels les plus importants et fournir un ensemble d'hypothèses fortes pour des processus de haut niveau d'interprétation.

Cette méthode procède en trois niveaux d'organisation. Un niveau de *groupements élémentaires* organise d'abord les éléments de contours en chaînes perceptuellement importantes. Le rôle de cette étape est de réduire la complexité des niveaux suivants en estimant de manière robuste l'importance structurelle des éléments de contours. Le niveau suivant recherche des *groupements intermédiaires* tels que des segments, des courbes simples ou des points d'intérêt. Ces éléments de représentation sont extraits à partir des chaînes groupées précédemment. Finalement, ceux-ci sont organisés selon des *groupements de haut niveau* dans une application de mise en correspondance structurelle de jonctions entre deux images.

Nos principales contributions sont l'apport d'un nouveau formalisme générique pour la construction et l'optimisation de réseaux d'éléments visuels localement connectés, l'utilisation de ce type de réseau pour extraire les contours les plus réguliers, une méthodologie de groupement hiérarchique de ces éléments visuels en hypothèses de structures plus complexes, et enfin, l'utilisation de ces structures pour la mise en correspondance de jonctions entre deux images.

Ce mémoire de thèse est divisé en deux parties. La première, théorique, est consacrée à la perception visuelle et au rôle que joue le groupement perceptuel en vision par ordinateur. Cette partie couvre l'ensemble des problèmes, théoriques et pratiques, que posent l'utilisation des contours pour l'analyse d'image.

– *Chapitre 1 : Vision naturelle et artificielle*

Nous posons dans ce chapitre la problématique de la vision par ordinateur. Afin de situer le contexte de notre travail, nous présentons les principales méthodologies appliquées en vision artificielle, en commençant par les différentes théories de la perception visuelle dont elles sont issues.

– *Chapitre 2 : Interprétation des scènes de contours*

Après avoir défini le contexte de notre travail, nous abordons dans ce chapitre la question des différentes représentations de l'environnement visuel. Le choix des contours comme élément de base de représentation est ensuite justifié. La problématique de la détection et la modélisation des contours, ainsi que leur rôle en vision par ordinateur sont abordés en détail dans ce chapitre.

– *Chapitre 3 : Groupement perceptuel en vision par ordinateur*

Ce chapitre est consacré à la place du groupement perceptuel en vision par ordinateur. En particulier, ce chapitre passe en revue les principes d'organisation perceptuelle issus de la psychologie de la vision et la manière dont ces principes ont été appliqués à la vision par ordinateur. Enfin, en guise de conclusion à cette première partie, nous exposons les principes de notre approche en regard de ces problèmes et des travaux antérieurs.

La seconde partie de ce mémoire concerne le travail effectivement réalisé au cours de cette thèse. Elle est divisée en trois chapitres, correspondant aux trois niveaux d'organisation perceptuelle de notre méthode.

– *Chapitre 4 : Saillance structurelle et groupements élémentaires*

Dans ce chapitre, nous abordons le premier niveau de groupement de notre approche. Il s'applique dès la détection de contours. Sa fonction est de faire ressortir les structures linéaires importantes présentes dans l'image de contours et de produire un ensemble de chaînes, ou groupements, correspondant à ces structures. Ces groupements servent de point de départ à l'extraction d'éléments de représentation réalisée aux niveaux suivants, en réduisant ainsi la complexité de la recherche visuelle aux seules structures d'intérêt de l'image.

– *Chapitre 5 : Eléments de représentation et groupements intermédiaires*

Nous montrons dans ce chapitre comment extraire de ces groupements élémentaires, un ensemble d'hypothèses géométriques utiles aux niveaux supérieurs de traitement. Dans un premier temps, nous présentons les principes de notre méthode de structuration hiérarchique, issus de l'analyse des groupements élémentaires. Cette méthode est ensuite détaillée pour chaque type d'hypothèse géométrique envisagée et illustrée par une application à différentes scènes réelles.

– *Chapitre 6 : Mise en correspondance structurelle*

Les éléments visuels extraits à partir des groupements précédents peuvent être comparés à un croquis sommaire des formes saillantes de la scène. Nous montrons dans ce chapitre comment établir des relations structurelles plus complexes à partir de ces éléments visuels afin de mettre en correspondance des structures issues de deux scènes. Nous illustrons enfin ce dernier niveau de groupement par une méthode de mise en correspondance de jonctions.

En conclusion, nous récapitulons enfin les principales contributions de notre approche, les améliorations à apporter et les perspectives ouvertes par cette thèse.