

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université de Tébessa  
Faculté des sciences exactes et des sciences de la nature et de la vie  
Département de mathématiques et d'informatique

# MÉMOIRE DE MAGISTER

En Informatique

**Option:** Ingénierie des Systèmes Informatiques (ISI)  
École Doctorale de l'ESI (EX INI)

Présenté par  
**ZAHROURI Ahmed**

Thème

## Une approche bayésienne à base d'agents pour la reconnaissance d'objets visuels

Devant le jury de soutenance composé de :

Mr. M. BENMOHAMED	<i>Professeur, Université de Constantine</i>	Président
Mr. S. CHIKHI	<i>Professeur, Université de Constantine</i>	Examineur
Mr. M.K. KHOLLADI	<i>Maître de conférences, Université de Constantin</i>	Examineur
Mr. S. MAZOUZI	<i>Maître de conférences, Université de Skikda</i>	Rapporteur

Année Universitaire : 2011/2012

---

# Remerciements

Je tiens à exprimer mes remerciements et ma vive gratitude :

A Monsieur Smaine Mazouzi Maître de Conférences à l'université de Skikda, qui a accepté d'être l'encadreur de ce mémoire, je le remercie tout particulièrement pour l'attention et le temps qu'il nous a consacré.

A Monsieur Mohamed Benmohamed (Professeur à l'université de Constantine), pour m'avoir fait l'honneur de présider le jury de soutenance. Qu'il trouve ici l'expression de ma sincère gratitude.

Aux membres du jury :

Monsieur Salim Chikhi (Professeur à l'université de Constantine), Monsieur Mohamed Khireddine Kholadi (Maître de Conférences à l'université de Constantine) pour avoir accepté de juger mon travail, et d'y être membres du jury de soutenance.

A tous les doctorants de l'ESI pour l'ambiance agréable qu'ils ont su créer.

A toute personne ayant participé de près ou de loin à l'élaboration de ce mémoire.

A ma famille et en particulier ma mère qui m'a donné l'éducation sans laquelle je n'en serais pas là aujourd'hui.

---

# Résumé

La vision artificielle cherche de nouvelles approches pour rendre la reconnaissance d'objets plus performante. Parmi les différentes méthodes proposées, les système multi-agents, et le framework bayésiens pour l'inférence de la meilleure interprétation du contenu d'une image.

Le présent travail de thèse propose une approche bayésienne à base d'agents pour la reconnaissance d'objets polyédriques, à partir d'images de profondeur. Dans un premier temps une population d'agents dits *agents détecteurs* repartis dans l'image pour suivre les points de contour issus d'une segmentation collective d'images de profondeur, afin de former des segments de droite et les enregistrer dans une structure de données. Cette dernière a été exploitée par un autre agent nommé *agent bayésien* pour inférer une meilleure représentation des objets.

Du fait des incertitudes dans les mesures et les observations, nous avons utilisé une approche probabiliste reposant sur l'utilisation de réseaux bayésiens. Ces réseaux permettent d'émettre des hypothèses sur l'existence et la localisation des segments d'objets, puis confirmer ou infirmer cette hypothèse. Enfin, en fonction du résultat de l'étape de vérification, de compléter le modèle 3-D de la scène.

**Mots clés** : Reconnaissance d'objets , Système Multi-Agents, Réseaux Bayésiens, Image de profondeur, Reconstruction 3D.

---

# Table des matières

<b>Table des figures</b>	<b>vi</b>
<b>Liste des tableaux</b>	<b>viii</b>
<b>Introduction Générale</b>	<b>ix</b>
<b>I La vision par ordinateurs</b>	<b>1</b>
1 Introduction . . . . .	1
2 La vision, une analyse . . . . .	1
3 Du signal au symbole . . . . .	2
4 La vision basée modèle . . . . .	6
5 Images de profondeur . . . . .	7
6 Conclusion . . . . .	8
<b>II Approches multi-agents en vision par ordinateur</b>	<b>9</b>
1 Introduction . . . . .	9
2 Les systèmes multi-agents . . . . .	9
2.1 Définitions . . . . .	10
2.2 Classification . . . . .	11
3 Systèmes multi-agents pour l'analyse et l'interprétation d'images . . . . .	12
3.1 Segmentation adaptative d'images par agents distribués . . . . .	12
3.2 Un système immuno-multi-agents pour le traitement d'images bio- logiques . . . . .	15
3.3 Surveillance multi-agents de scènes dynamiques . . . . .	17

3.4	Un système co-évolutionnaire pour la segmentation d'images . . . . .	19
4	Conclusion . . . . .	21
<b>III Approches Bayésiennes en vision par ordinateur</b>		<b>22</b>
1	Introduction . . . . .	22
2	Réseaux Bayésiens . . . . .	22
2.1	Probabilité conditionnelle . . . . .	22
2.2	Théorème de Bayes . . . . .	23
2.3	Réseaux bayésiens . . . . .	23
2.4	La notion de causalité . . . . .	24
2.5	Les inférences . . . . .	24
2.6	Apprentissage . . . . .	25
3	Les applications . . . . .	25
3.1	Domaines d'applications . . . . .	25
4	Réseaux Bayésiens en vision par ordinateur . . . . .	26
4.1	Reconnaitre des modèles géométriques simples . . . . .	26
4.2	Le contrôle de la perception sélective . . . . .	30
4.3	Surveillance du trafic routier . . . . .	31
5	Conclusion . . . . .	32
<b>IV Une approche bayésienne distribuée pour la reconnaissance d'objets</b>		<b>33</b>
1	Introduction . . . . .	33
2	Description de l'approche . . . . .	33
2.1	Agents détecteurs(Agents Père, Agents Fils) . . . . .	34
2.2	Agent bayésien . . . . .	41
2.2.1	Prédiction : génération des hypothèses . . . . .	42
2.2.2	Vérification des hypathèses émises . . . . .	45
2.3	Les agents reconnaisseurs . . . . .	47
2.3.1	Les agents reconnaisseurs : . . . . .	47
2.3.2	L'agent coordonnateur : . . . . .	48
3	Conclusion . . . . .	49
<b>V Résultats expérimentaux</b>		<b>50</b>
1	Introduction . . . . .	50
2	L'environnement de l'expérimentation . . . . .	50
2.1	Le langage de programmation . . . . .	50

## Table des matières

---

2.2	Modélisation Bayésienne(BayServer) . . . . .	51
3	La description des agents du système . . . . .	51
3.1	Agent détecteur . . . . .	51
3.2	Agent Bayésien : . . . . .	54
4	Conclusion . . . . .	57
	<b>Conclusion générale</b>	<b>58</b>
	<b>Références bibliographiques</b>	<b>60</b>

---

## Table des figures

I.1	Exemple d'interprétation différente pour un même symbole . . . . .	2
I.2	La segmentation, reconstruction et la reconnaissance . . . . .	3
I.3	Complémentarité des approches ascendantes et descendantes . . . . .	5
I.4	Principe de la vision basée modèle . . . . .	7
I.5	Principe de télémétrie par triangulation . . . . .	8
II.1	Région de travail de rayon $R$ d'un agent en $(i, j)$ . . . . .	13
II.2	(a) La reproduction; (b) La diffusion . . . . .	14
II.3	Agents réactifs pour la détection d'anneaux concentriques. (a) Plaice otolith de 8 ans; (b) Architecture BRIC d'un agent . . . . .	16
II.4	Capteurs d'un agent . . . . .	16
II.5	Réseau bayésien des agents utilisés : (a) Agent de comportement; (b) Agent de situation. . . . .	18
II.6	Co-évolution des tributs. (a) Identités initiales des agents; (b) Echange d'identités entre agents; (c) Influence mutuelle entre agents; (d) Tribus stables et homogènes . . . . .	20
III.1	Un réseau bayésien . . . . .	23
III.2	Exemple de causalité . . . . .	24
III.3	Discrétisation de l'orientation des segments et exemple de coins. . . . .	26
III.4	Motifs élémentaires : coins à jonction double. . . . .	27
III.5	Réseau hypothèse . . . . .	28
III.6	Vérification d'hypothèse . . . . .	29

III.7 Modèle d'une fenêtre . . . . .	30
IV.1 Schéma de l'approche proposée . . . . .	34
IV.2 Les positions des agents sur les points de contour . . . . .	34
IV.3 Les deux types de points :(a) Point de coin; (b) Point isolé. . . . .	36
IV.4 Comportement de l'agent père . . . . .	37
IV.5 Comportement de l'agent fils . . . . .	38
IV.6 Réseau de génération d'hypothèses . . . . .	43
IV.7 Classe élémentaires et hypothèses associées :(a) Extrémités proches; (b) Extrémités éloignés. . . . .	44
IV.8 Réseau de vérification d'hypothèses . . . . .	46
IV.9 Génération d'une hypothèse de correspondance objet-modèle . . . . .	47
IV.10Hypothèses de correspondance objet-modèle (a,b); (c) Modèles d'objets . .	49
V.1 Les diagrammes : (a) Diagramme de classes; (b) Diagramme de séquences.	52
V.2 Image de synthèse : (a) Une image représente un cube; (b) Position d'agents sur le cube . . . . .	53
V.3 Nombre de segments détectés selon le nombre d'agents lancés (NBagent,NBseg) : (a) (NBagent=2; NBseg=2); (b) (NBagent=5; NBseg=5); (c) (NBagent=10; NBseg=9); (d) (NBagent=26; NBseg=10). . . . .	54
V.4 Image de synthèse : (a) Une image représente un cube; (b) Un cube présente des discontinuités marquées en rouge . . . . .	55
V.5 Reconstruction de l'image : (a) Segments détectés; (b) Image acquise pré- sédement; (c) Reconstruction de l'image avec l'ajout d'un segment $S_{15}$ . . .	56

---

## Liste des tableaux

IV.1	Table de probabilités conditionnelles $p(\textit{hypoth\grave{e}se} \textit{classe})$ pour les extrémités proches . . . . .	45
IV.2	Table de probabilités conditionnelles $p(\textit{hypoth\grave{e}se} \textit{classe})$ pour les extrémités éloignées . . . . .	45
V.1	Le nombre de segments détectés selon le nombre d’agents lancés (Figure V.1)	53
V.2	Segments trouvés et leurs positions . . . . .	54

---

# Introduction Générale

La vision est sans doute, le sens le plus important donné à l'homme ; il lui permet de saisir, en un temps très court, un volume très considérable d'informations concernant son environnement externe. En effet, c'est le système de vision qui permet à l'être humain d'agir sur cet environnement avec une haute performance. Il lui permet aussi de mieux communiquer avec les autres avec rapidité et efficacité.

Les robots, au départ, non sensibles à leurs environnements, sont devenus aujourd'hui, capables de capter l'information sur cet environnement et de réagir en conséquence, à des événements externes imprévus. Cette sensation est principalement développée par la vision artificielle et plus précisément la vision robotique qui est due aux progrès de l'électronique et de l'informatique. Des capteurs visuels très sophistiqués sont maintenant utilisés. Ils permettent d'acquérir, d'une manière directe ou indirecte, des images tridimensionnelles d'une scène réelle.

La vision artificielle est un domaine de recherche extrêmement difficile, principalement parce qu'il oblige le chercheur à essayer d'imaginer les mécanismes impliqués dans la vision. Ainsi, il ne peut ignorer les découvertes faites par les chercheurs travaillant dans les sciences de la vie. Pour cela, rien n'empêche le chercheur en vision artificielle, de copier le système de vision naturelle, lorsqu'il s'agit de résoudre un problème de robotique. De ce fait, la vision artificielle pourrait être considérée à la fois comme une science fondamentale et une science appliquée.

Un système de vision artificielle est conçu pour analyser une image afin de produire une description du contenu de celle-ci. A partir de laquelle, un raisonnement pourra être initié pour accomplir la tâche demandée. Elle doit donc extraire certaines propriétés phy-

---

siques des objets figurants dans l'image. Actuellement, la vision artificielle est utilisée dans plusieurs domaines tels que l'inspection automatique des chaînes de fabrication, l'autonomie d'un robot dans un environnement réel, l'analyse et l'interprétation de l'imagerie en médecine, l'aérospatial et le multimédia.

Pour analyser et interpréter une image, le système de vision artificielle procède principalement en trois étapes. Dans la première étape, qui correspond au traitement de bas niveau, l'image est segmentée pour en extraire les indices visuels pertinents. La deuxième étape, qui correspond au niveau intermédiaire, consiste en la structuration des indices visuels en une représentation symbolique décrivant d'une manière non ambiguë la sémantique des objets à reconnaître. Enfin, la troisième étape correspond au traitement de haut niveau et consiste à faire un raisonnement sur la description symbolique de la scène pour permettre le matching objet-modèle.

Avec un environnement sans modèle ou mal modélisé, tel qu'une image, la coopération d'entités distribuées opérant d'une façon autonome peut s'avérer une solution efficace, pour laquelle un processus de reconnaissance continue à se dérouler correctement dans les situations où il devrait s'arrêter ou produire de faux résultats. Par ailleurs, l'incertitude de l'information visuelle, s'apprête bien à l'utilisation de frameworks bayésiens pour l'inférence de la meilleure modélisation du contenu d'une image.

Ansı, nous proposons dans ce mémoire, une approche bayésienne à base d'agents pour la reconnaissance de scènes composées d'objets polyédriques à partir d'images de profondeur. La reconnaissance se déroule en trois étapes :

- *Détection* : Dans ce niveau nous exploitons une population d'agents dits *agents détecteurs* évoluant sur l'image en suivant les points de contour dont le but de former des segments de droites, et les enregistrer dans une structure de données nommée appropriée.
- *Reconstruction* : Un autre agent nommé agent bayésien opère dans ce niveau de notre système, et qui exploite les connaissances *a priori* sur les scènes polyédriques (ces connaissances sont codées sur des réseaux bayésiens), afin de former des objets complexes à partir de segments détectés.
- *Reconnaissance* : Dans ce niveau un agent reconnaisseur qui prend en charge la tâche de reconnaissance et engendre le processus d'identification de type Objet-Modèle.

---

Ce mémoire est organisé de la façon suivante :

Dans le chapitre 1, nous présentons une introduction sur notre domaine de recherche, qui est la vision par ordinateur. On montre à la fois, le modèle de David Marr et ses trois étapes de traitement : La segmentation, la reconstruction, et la reconnaissance.

Dans le chapitre 2, nous analysons quelques contributions typiques en vision par les systèmes multi-agents, en montrant les spécificités de ces contributions, ainsi que leurs limitations majeures. Nous focalisons la discussion sur les aspects multi-agents et comment sont ils utilisés pour surmonter les problèmes liés à l'analyse et l'interprétation des images.

Le chapitre 3 est consacré aux réseaux bayésiens et leur contribution en vision par ordinateur, nous commençons ce chapitre par un rappel sur ces derniers, et ensuite, nous exposons quelques travaux qui utilisent les réseaux bayésiens pour traiter les problèmes de la vision par ordinateur.

Nous présentons dans le chapitre 4 une approche bayésienne à base d'agents pour la reconnaissance d'objets. Dans ce chapitre, nous essayons de détailler notre approche qui est constituée de trois niveaux, détection de segments de droites, reconstruction bayésienne et enfin la reconnaissance.

Le chapitre 5 est réservé aux résultats expérimentaux. Nous exposons dans ce chapitre, la plateforme de développement, et nous présentons et commentons quelques résultats obtenus par notre système.

Nous terminons ce mémoire par une conclusion générale tout en mentionnant les extensions et perspectives futures de ce travail.

---

---

# Chapitre I

---

## La vision par ordinateurs

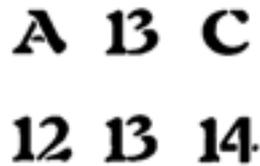
### 1 Introduction

le but de la vision artificielle est de concevoir un système capable de traiter les informations visuelles de façon aussi versatile et sophistiquée que l'être humain. De façon plus concret, le désir est de développer des moyens qui permettraient à une machine d'interagir avec l'environnement en se basant sur ses capteurs visuels. Lorsque qu'on vous demande de décrire ce que vous voyez autour de vous, vous n'avez aucune difficulté à distinguer tous les objets qui vous entourent. Si nous nous attardons sur l'information qui est réellement disponible au niveau de la rétine, nous nous apercevons que c'est simplement un ensemble de points lumineux. Pour chacun de ces points nous avons une indication sur l'intensité lumineuse et la couleur. Bien qu'il y ait déjà un certain nombre de traitements, le cerveau, sur la base des percepts reçus, construit des concepts appuyés sur toutes les connaissances acquises par notre expérience. Souvent, il n'y a aucune correspondance directe entre l'information sensorielle telle que la lumière, la couleur et la représentation finale (les objets 3-D). C'est donc au système de vision de fournir ces connaissances nécessaires à l'interprétation [Nat96].

### 2 La vision, une analyse

Les mécanismes de reconnaissance ne sont pas encore complètement compris. Il est admis que l'humain interprète les images selon leur contexte. Par exemple l'image de la figure I.1 nous montre à la première ligne, une séquence de lettres : A, B, C. La séquence de la deuxième ligne présente les chiffres 12, 13, 14. Nous pouvons constater que la lettre B et le nombre 13 ont la même forme. Quelle indication a permis au cerveau d'interpréter ce symbole une fois en B et l'autre en 13 ? [Nat96]. Cet exemple permet de voir qu'il n'est

pas possible de définir une entité sans se référer à son contexte.



**Figure I.1** – Exemple d'interprétation différente pour un même symbole

La vision artificielle est fondamentalement différente du traitement du signal visuel. Ce dernier étudie les transformations de signal, n'extrayant pas explicitement les informations symboliques du signal [BB82]. Les systèmes automatiques ne sont pas capables de travailler au niveau de concepts. Pour la vision artificielle, la manipulation des images est étroitement liée à leur interprétation [Nat96].

### 3 Du signal au symbole

Les êtres humains perçoivent l'environnement comme étant séparé en entités physiques. Nous pouvons poser plusieurs questions sur ce point [Nat96] :

- Quel genre de processus permet de passer de la rétine (bas niveau) au concept (haut niveau) ?
- Quelle description symbolique est la plus appropriée ?
- Comment l'information se propage ?
- Quelle hiérarchie est utilisée dans ce processus ?

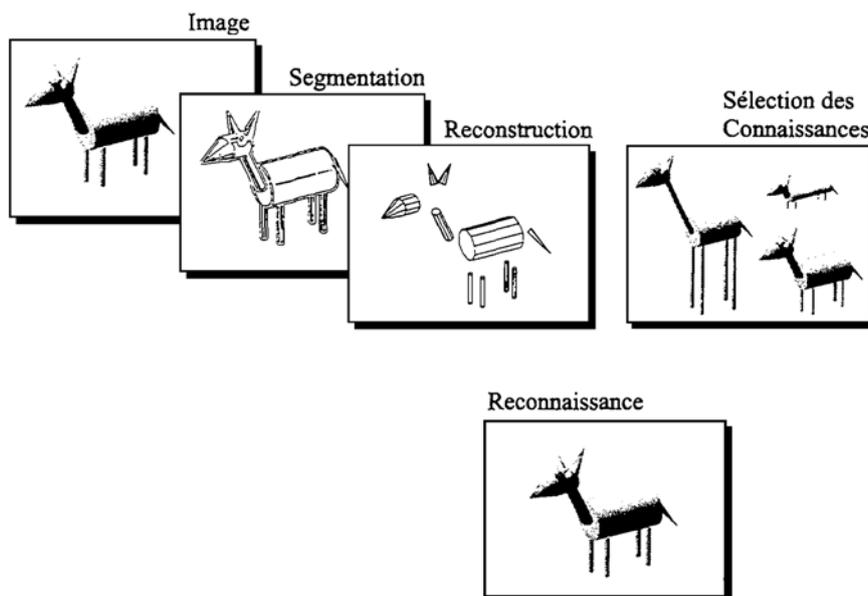
Le processus permettant d'extraire le concept du percept, est appelé inférence inductive du fait qu'une description du monde est inférer des images.

La deuxième question relate de l'immense choix de représentations possibles. Il est admis que l'être humain est capable de reconnaître plus de cent mille objets différents [Tso88], alors que les principes et mécanismes de mémorisation restent encore inconnus.

David Marr a proposé, vers les années 70, un modèle décrivant le traitement et la représentation de l'information visuelle. Il soulignait le fait que le monde est constitué d'objets

solide non-déformables qui occupent des points de l'espace différents. Les trois modes de représentation de la vision humaine, selon Marr, sont : La *première esquisse* (primal sketch) 2-D définie par les changements d'intensité de l'image selon sa géométrie. En parlant d'une ou plusieurs images un processus d'extraction de caractéristiques produit une description en termes d'attributs bidimensionnels ; c'est l'ensemble de contours d'objets, leur texture et leur orientation. Ensuite, nous avons l'*esquisse 2.5-D*, qui est une description centrée sur l'observateur (description dépendante de la position de l'observateur), et qui constitue l'entrée d'un processus qui calcule des propriétés tri-dimensionnelles locales, relatives à la scène tel que les contours, l'éloignement de surfaces visibles et leur normales. L'analyse des contours ainsi que des textures nous informe sur la géométrie de l'image. La décomposition des ombrages nous renseigne sur la photométrie et finalement, l'étude de reflets nous informe sur la colorimétrie. Signalons que des techniques telles que l'analyse du mouvement par décomposition de séquences d'images et la stéréoscopie peuvent être utilisées par l'esquisse 2.5-D. Enfin, le modèle 3-D, est la représentation la plus complète, centrée sur la scène, permettant l'observation sous n'importe quel angle. Ce dernier décrit les objets dans le but de les reconnaître et de les manipuler [Nat96].

Le paradigme de Marr peut être traduit par trois étapes de traitement, la *segmentation*, la *reconstruction*, et la *reconnaissance*.



**Figure I.2** – La segmentation, reconstruction et la reconnaissance

La *segmentation* est l'opération qui consiste à subdiviser une scène réelle en ses parties constituantes ou objets. Il en résulte un découpage de l'image en régions significatives. En d'autres termes, la segmentation est la recherche de primitives de surfaces et de contours. A ce point, les objets sont caractérisés par une séquence de pixels qui appartiennent soit à leur contour fermé, soit à leur surface. Pour aboutir à une reconnaissance d'objets, nous sommes obligés d'étiqueter chaque forme segmentée. Cette étape est appelée *reconstruction* et nous permettra la différenciation des objets présents dans la scène. En les regroupant par éléments appartenant à une même primitive. La prochaine étape est la *reconnaissance* proprement dite des objets. Nous utilisons notre *connaissance* des objets pouvant être mis en correspondance avec les primitives extraites. Par exemple, dans la figure I.2, nous pouvons voir que grâce à l'étape de reconstruction, nous avons effectué une sélection sur les connaissances, en ne laissant que les objets contenant les primitives extraites lors de la segmentation. La dernière opération nécessaire à la compréhension de la scène, est *l'interprétation* qui donne un sens à l'ensemble des objets trouvés. Pour la reconnaissance, nous pouvons dénombrer différentes techniques. La reconnaissance directe superpose un objet connu à un objet inconnu et on la différencie au niveau des modèles. Une autre méthode s'appuie sur les notions de distances euclidiennes entre la silhouette et le descripteur [Nat96].

Concernant le flux des données, deux types de stratégies sont mises en jeu : *ascendante* (bottom-up) et la *descendante* (top-down) (Figure I.3). Les stratégies descendantes tentent, en partant d'une modélisation de la scène, c'est-à-dire d'une connaissance (géométrique, photométrique ...) des objets, de reconnaître dans une image si tout ou une partie de ces modèles sont présents. Nous procédons des modèles à l'image. Même en connaissant *a priori* tous les objets présents dans la scène, nous ne pouvons ni prédire leur position et orientation, ni savoir s'ils se superposent mutuellement. Dans l'approche *ascendante*, on cherche d'abord à regrouper des régions de l'image selon des attributs communs et ensuite à modéliser ces régions. Nous allons de l'image aux modèles. Finalement, en extrapolant ces modèles, elle assure la reconnaissance de l'objet [Nat96].

Résumons le problème du flux des données par deux questions-clefs :

- Bottom-up : Est-ce que l'information contenue dans le signal est suffisante pour construire un modèle sans créer trop de descriptions ambiguës ?
- Top-down : Est-il possible de projeter des modèles directement sur les données sans pour autant aboutir à une explosion combinatoire de mise en correspondances ?

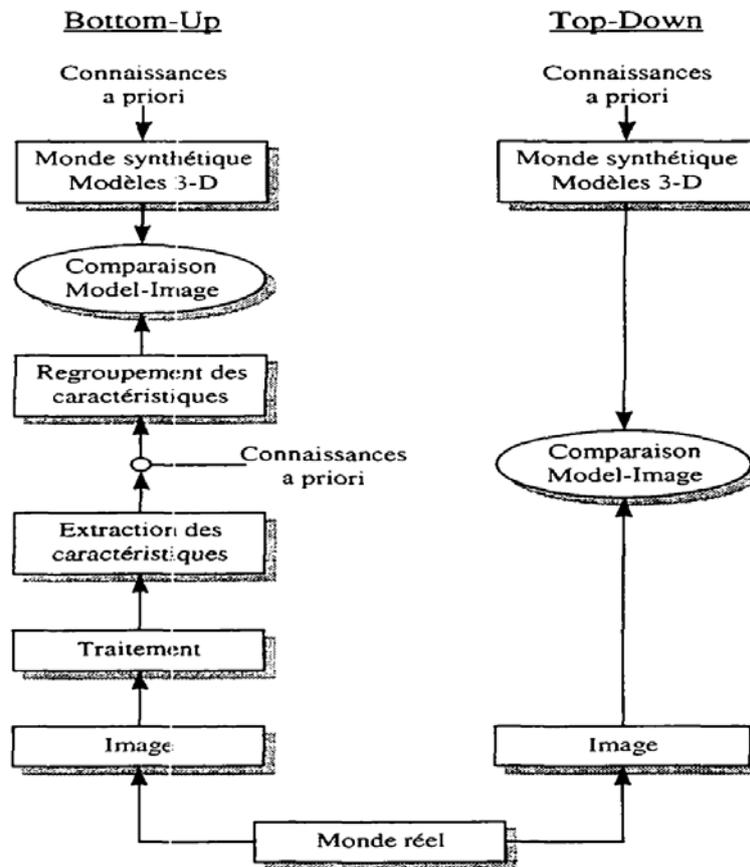


Figure I.3 – Complémentarité des approches ascendantes et descendantes

La première approche descendante a l'avantage de connaître exactement les objets, ce qui permet d'adapter les techniques de recherche, selon la tâche à accomplir. Ce qui veut dire que si on cherche, par exemple, une boîte rouge sur une table blanche, la seule information de couleur suffira à la reconnaissance. Le grand désavantage de cette technique est qu'il faut un modèle initial pertinent [Nat96].

L'approche ascendante présente l'avantage de pouvoir utiliser une large palette d'outils (extraction de contours, regroupement de segments, arbres binaires,...) qui ont prouvé leur efficacité depuis longtemps. La principale difficulté que présente cette technique est la nécessité de faire appel à une certaine "intelligence" à plusieurs étapes du traitement et regroupement. En effet, pour regrouper des segments afin de reconstituer une surface, il faudra, par exemple, identifier uniquement les arrêtes d'une surface et les distinguer des artefacts qui apparaissent aux frontières des ombres ou des textures. Ces techniques de

regroupement d'attributs pour reconstituer les modèles sont très sensibles au bruit. Le processus étant directement dépendant des images (données brutes), toute erreur sur ces dernières, entraînerait des dégradations du taux de reconnaissance [Nat96].

Le problème consiste à choisir quelle est la technique permettant de constituer un système de vision performant et insensible au bruit. Il est plus naturel de chercher un objet en se basant sur son modèle, mais ce modèle doit pouvoir se rajuster (taille, orientation, position) d'après la scène présentée en utilisant des informations bas niveau [Nat96].

## 4 La vision basée modèle

Le but d'un système de vision basé modèle est de reconnaître des objets dans une scène par leur mise en correspondance avec des modèles mémorisés. Cette tâche peut s'avérer très complexe si les objets sont partiellement cachés ou se chevauchent. Généralement, un tel système de reconnaissance dispose, dans une base de données appelée *modèles*, d'un ensemble d'attributs caractéristiques de l'objet à rechercher. La reconnaissance implique une reconnaissance d'attributs spécifiques dans l'image et une mise en correspondance avec les modèles à disposition. De ce fait, seuls les objets dont il existe des modèles pourront être reconnus [Nat96].

Les systèmes de reconnaissance d'objets basés modèles sont pour la plupart des systèmes de prédiction-vérification et dont le fonctionnement peut être représenté par le schéma de la figure I.4. Nous pouvons reconnaître les trois étapes clef d'un tel système. Notamment la création du modèle, l'extraction d'attributs et la reconnaissance. Dans une première phase, il faut choisir les attributs qui seront retenus de manière à caractériser chaque modèle de la façon la plus efficace. Généralement, un objet est caractérisé par sa géométrie, sa texture, sa taille, ou tout autre attribut spécifique [Nat96].

La seconde étape consiste à extraire les attributs de la scène pour ensuite les mettre en correspondance avec ceux des modèles mémorisés. Cette étape est très dépendante de la tâche de recherche et est souvent guidée par les *connaissances a priori* de la scène observée [Nat96].

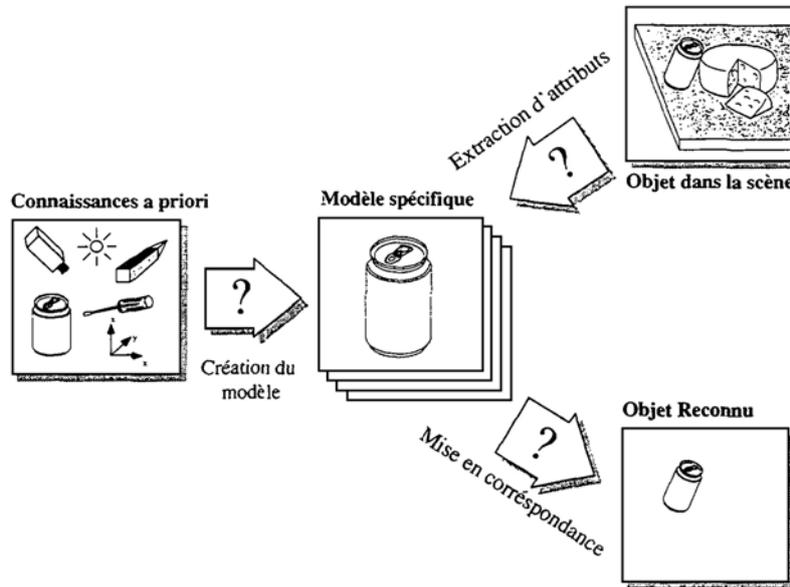


Figure I.4 – Principe de la vision basée modèle

## 5 Images de profondeur

Une image de profondeur est une image 3D où à chaque pixel  $(x, y)$  correspond la distance  $Z(x, y)$  entre le plan du capteur télémétrique et le point correspondant de la scène. Dans une telle image, les régions représentent les parties visibles des surfaces des différents objets figurant dans la scène. Selon la nature de données capturées, les techniques d'acquisition peuvent être classées en deux catégories distinctes [AA93] :

- Les techniques actives, pour lesquelles un faisceau laser ou ultrasonore est utilisé. Celles-ci sont capables de mesurer directement la profondeur des objets perçus. Par conséquent, les données visuelles sont directement tridimensionnelles ;
- Les techniques passives, pour lesquelles l'énergie ambiante émise ou réfléchiée, par les objets de la scène, est utilisée pour déduire les distances 3D de ces objets.

Les systèmes basés sur la triangulation sont des systèmes actifs utilisant la loi des sinus [HM95] pour mesurer la distance vers un point de la scène (Figure I.5). La distance  $h$ , d'un point de l'objet, au plan de balayage est calculée comme suit :

$$h = d \frac{\tan(A) \tan(B)}{\tan(A) + \tan(B)} \quad (\text{I.1})$$

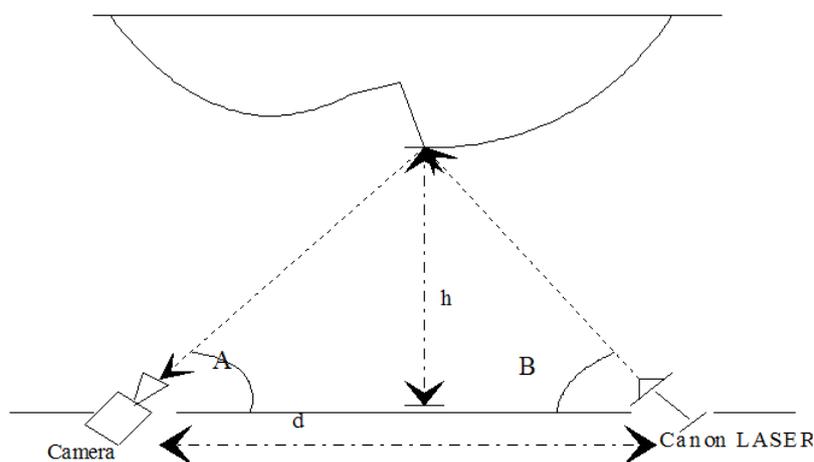


Figure I.5 – Principe de télémétrie par triangulation

L'acquisition consiste en l'échantillonnage spatial d'une scène observée par un faisceau de lumière concentré (généralement un faisceau laser) dont l'angle d'émission est préalablement connu. Le faisceau diffusé en chaque point éclairé de la scène, est focalisé par un objectif sur un photodétecteur (caméra), et ce dans l'objectif de mesurer l'angle de réception du faisceau réfléchi sur le point éclairé. La connaissance de l'angle d'émission et de l'angle de réception permet, par triangulation, de calculer la profondeur  $Z(x, y)$  de chaque point  $(x, y)$  de la scène. Après échantillonnage de toute la scène, le système génère un tampon  $Z$  représentant l'image de profondeur de la scène observée.

## 6 Conclusion

Dans ce chapitre, nous avons introduit notre domaine de recherche qui est la vision par ordinateur. On montre à la fois, le modèle de David Marr et leurs trois étapes de traitement : la segmentation, la reconstruction, et la reconnaissance. Nous avons montré aussi dans ce chapitre les techniques de vision basées modèles, qui y seront la base de notre approche dans le chapitre IV.

---

---

# Chapitre II

---

## Approches multi-agents en vision par ordinateur

### 1 Introduction

Les systèmes multi-agents pour la résolution de problèmes en vision par ordinateur héritent les avantages de l'approche agent, notamment la résolution d'un problème d'une manière collective et distribuée. Dans de tels systèmes, chaque agent à lui-même, par manque de compétence générale ou par manque de données globales, ne peut pas résoudre un problème dans son intégralité. Les agents sont donc appelés à interagir et à coopérer afin de combler leur insuffisance personnelle, en terme de compétence et de perception. L'évolution de la dynamique du système, qui résulte de l'interaction des agents entre eux et de leur action sur leur environnement, permet de faire émerger une solution collective au problème. Avant de procéder à l'analyse des systèmes multi-agents en vision par ordinateur, les plus référenciés dans la littérature, nous donnons un aperçu général sur les systèmes multi-agents.

### 2 Les systèmes multi-agents

Dans les systèmes multi-agents, un agent est communément considéré comme une entité software qui, contrairement à un objet, exhibe les caractéristiques suivantes [WJ95] [Woo02] :

- L'agent est autonome, et possède son propre thread d'exécution. Il n'est par conséquent soumis à aucun contrôle centralisé d'exécution.

- Il est situé dans un environnement particulier qui peut sentir et sur lequel il peut agir.
- Un agent est proactif, et n'agit pas uniquement en réponse aux événements en provenance de son environnement. Il exhibe un comportement dirigé par un but, où l'agent est en mesure de prendre de l'initiative [ZJW03].
- Un agent est sociable car il interagit avec d'autres agents, directement en utilisant un langage de communication d'agents Agent Communication Language : ACL [GK94], ou indirectement via l'environnement.

### 2.1 Définitions

Afin de bien montrer la spécificité d'un agent, comparé à un programme informatique ordinaire, nous donnons plusieurs définitions, citées dans des travaux de références en agents et systèmes multi-agents. Le concept qui revient avec la plupart de ces définitions est l'autonomie.

- Selon Shoham [Sho93], Souvent, quand les gens utilise le terme agent ils se réfèrent à une entité qui fonctionne d'une manière continue et autonome, et ce dans un environnement dans lequel d'autres processus peuvent avoir lieu, et d'autres agents existent.
- Russell stipule qu'un agent est une entité qui sent et qui agit sur son environnement [Rus97].
- Selon la définition de l'agent 'MuBot', le terme 'agent' est utilisé pour représenter deux concepts orthogonaux. Le premier est la capacité de l'agent pour s'exécuter d'une manière autonome. Le second est la capacité de l'agent d'exécuter le raisonnement orienté domaine d'application.
- Selon IBM, les agents intelligents sont des entités logicielles qui effectuent un certain jeu d'opérations au profit d'un utilisateur ou un autre programme, avec un certain degré de liberté ou d'autonomie et ainsi, utilisent des connaissances ou des représentations des objectifs ou des désirs des utilisateurs.
- Selon Franklin et Gasser, un agent autonome est un système placé dans un environnement dont il fait partie. L'agent sent l'environnement et y agit, dans la poursuite de son propre but [FG96].

Suite aux définitions précédentes, nous pouvons constater que certains auteurs ont essayé de donner une définition de l'agent comme une entité isolée. Leurs travaux ont

principalement concerné les architectures internes des agents et leur capacité d'action ou de mobilité. Cependant, d'autres auteurs voient les agents comme entités agissant collectivement avec d'autres agents d'où le paradigme multi-agents [CBF03].

Formellement, Ferber a défini un système multi-agent comme un système caractérisé par les éléments suivants [Fer95] :

- Un espace métrique  $E$ , dit environnement ;
- Un ensemble d'objets  $O$ , incluant les objets passifs et les objets actifs (agents) ;
- Un ensemble d'agents  $A$ , sous ensemble de  $O$  ;
- Un ensemble de relations  $R$ , défini entre les objets de l'ensemble  $O$  ;
- Un ensemble d'opérations  $Op$  permettant aux agents d'agir sur les objets et sur l'environnement.

Selon ces éléments, dans un système multi-agents, l'agent qui peut être une entité matérielle ou logicielle dotée d'une capacité de perception et d'action, est placé dans l'environnement dans lequel il pourra à tout moment être situé. Afin d'interagir et de coopérer, les agents communiquent entre eux, et ce soit d'une manière explicite en utilisant l'information partagée ou l'envoi de message ; soit d'une manière implicite en utilisant des marques que l'on dépose dans l'environnement. Ainsi, les agents sont dotés d'un comportement autonome, conditionné par leur perception, leur connaissance, et leur compétence.

### 2.2 Classification

Suivant leur granularité et leur méthode de représentation des connaissances, les agents sont classés en deux catégories distinctes :

1. Les agents cognitifs sont caractérisés par une granularité forte. Ils utilisent un niveau élevé de représentation de connaissances ; d'où leur capacité de raisonnement et de négociation. Ces agents utilisent des langages de haut niveau, leur permettant d'échanger leurs connaissances et de coordonner explicitement leurs actions.
2. Les agents réactifs constituent l'alternatif aux agents cognitifs pour la modélisation des comportements intelligents dans les systèmes à base d'agents. Il s'agit d'entités simples, qui réagissent d'une manière réactive aux stimuli qui proviennent de leur environnement, en changeant de comportement et/ou en agissant sur les objets de

l'environnement. Ces agents ne disposent pas d'une représentation symbolique de leur monde et par conséquent ne disposent pas de mécanismes complexes de raisonnement. L'intelligence dans les systèmes multi-agents réactifs, consiste en une propriété émergente de l'activité entière du système. Souvent on s'inspire des comportements de larges sociétés d'êtres simples, telles que les insectes, les poissons et les oiseaux [Dro94].

### 3 Systèmes multi-agents pour l'analyse et l'interprétation d'images

La plupart des travaux, ayant proposé des systèmes multi-agents en vision par ordinateur, ont opté pour l'approche d'agents situés. Ce choix est motivé principalement, par le fait que l'information brute traitée en vision artificielle est locale, située au niveau des pixels de l'image. L'image représente l'environnement dans lequel les agents sont placés pour l'explorer à la recherche des différents objets qui représentent son contenu. Selon ce schéma, les agents qui, d'une manière ou d'une autre, parcourent l'image à la recherche de caractéristiques géométriques bien déterminées, interagissent et coopèrent pour inter-échanger mutuellement l'information qui leur manquait, et prendre des décisions concernant les sites de l'image analysée. Dans ce qui suit, de cette section, nous présentons, l'essentiel de quelques travaux, récents, utilisant l'approche multiagents pour l'analyse et l'interprétation d'images.

#### 3.1 Segmentation adaptative d'images par agents distribués

Un système multi-agents pour la segmentation d'images à résonance magnétique (*IRM*) du cerveau, était proposé par Liu et al. [LT99]. Dans cette approche, l'accent est mis sur la recherche par une population d'agents de segmentation travaillant de manière optimale. Chaque agent procède ici à l'étiquetage d'un point de l'image : il s'agit donc de générer la distribution d'agents la mieux adaptée à la distribution de points rencontrée dans l'image. La reproduction d'agents, fonction des variations locales de l'image, peut permettre de résoudre une des difficultés de la segmentation d'image : le fait que les régions d'une image peuvent être homogènes à des degrés différents. L'image est décomposée en segments homogènes. Chaque agent va alors chercher un segment homogène afin de s'y reproduire. Chaque segment ou région homogène est défini à l'aide de critères d'homogénéité.

### Description des critères d'homogénéité

Chaque point de l'image d'intensité  $I(i, j)$  est étiqueté en fonction de critères d'homogénéité dans sa zone de travail (Figure II.1) calculés à partir de trois index : contraste, moyenne et variance.

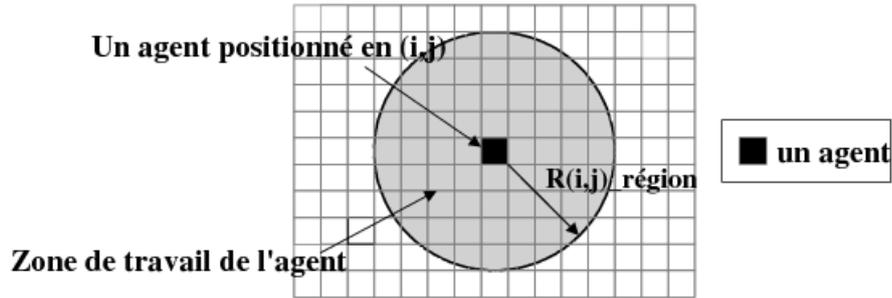


Figure II.1 – Région de travail de rayon  $R$  d'un agent en  $(i, j)$

Les critères sur ces trois index sont ainsi définis :

- le contraste :  $G(i, j)_{région} \in [g1, g2]$
- la moyenne :  $M(i, j)_{région} \in [m1, m2]$
- la variance :  $V(i, j)_{région} \in [v1, v2]$

où  $m_1$ ,  $m_2$ ,  $v_1$  et  $v_2$  sont prédéfinis et  $g_1$ ,  $g_2$  sont définis à partir de l'histogramme global sur l'image et du nombre de tissus à segmenter. Le contraste nous donne le nombre de pixels, situés à l'intérieur de la zone de travail, dont l'intensité est proche de celle de l'agent en  $(i, j)$ . Dans le cas où  $G(i, j)_{région}$  serait trop faible on va alors déplacer l'agent initialement en  $(i, j)$ . La variance nous permet d'avoir un aperçu de la façon dont les données sont réparties autour de la moyenne et ainsi d'optimiser le choix de la région de travail.

### Le processus de segmentation

Un nombre donné d'agents est tout d'abord placé de manière aléatoire dans l'image. A chaque agent est associée une zone de travail, un vecteur définissant des directions de reproduction. Chaque agent peut soit se reproduire soit chercher à se déplacer en fonction de sa réponse aux critères d'homogénéité calculés dans sa zone de travail. Si un agent détecte qu'un pixel de coordonnées  $(i, j)$  localisé dans sa zone de travail  $R(i, j)$ , satisfait

## II.3 Systèmes multi-agents pour l'analyse et l'interprétation d'images

les critères d'homogénéité alors il va se reproduire sous la forme d'un nombre fini de fils selon des directions spécifiques données par le vecteur de direction correspondant.

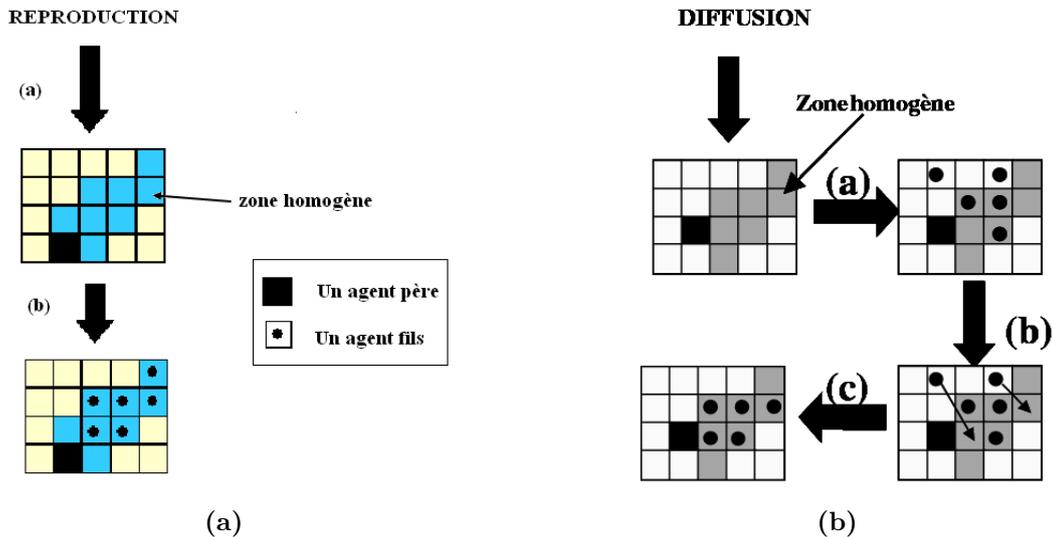


Figure II.2 – (a) La reproduction ; (b) La diffusion

### Reproduction :

- (a) : Un agent père étiquette un pixel se trouve sur zone homogène.
- (b) : Il va alors se reproduire dans des directions déduites du vecteur de reproduction.

### Diffusion :

- (a) : Un agent père placé dans une région homogène va pouvoir se reproduire suivant des directions obtenues par le vecteur de reproduction.
- (b) : Les fils n'appartenant pas à la région homogène vont commencer leur diffusion à l'intérieur de leur zone de travail suivant des directions héritées de leur parents et frères.
- (c) : Après la phase de diffusion certains fils se repositionnent dans la zone homogène.

Dans le cas où l'agent détecte qu'il se situe en dehors d'une région homogène, c'est-à-dire que les critères d'homogénéité ne sont pas vérifiés, il va alors chercher à « diffuser »

## II.3 Systèmes multi-agents pour l'analyse et l'interprétation d'images

---

en bougeant vers un autre pixel au sein de sa zone de travail. La direction de diffusion est déterminée à partir du vecteur de direction en fonction des positions de son père et de ceux de ses frères qui vérifient les critères d'homogénéité (Figure II.2). Quand un agent se trouve sur un pixel vérifiant tous les critères, alors il va étiqueter ce pixel pour signifier qu'il appartient bien à la zone homogène recherchée. Les directions de reproduction et de diffusion sont déterminées en fonction de deux vecteurs de reproduction et diffusion.

### Discussion

Cette approche nous propose un système qui exploitant à la fois peu de connaissances a priori et peu de contraintes : un tel système présente des capacités certaines d'adaptation aux variations locales de l'image. En revanche, il s'avère difficile à contrôler, et peut conduire à des erreurs de segmentation difficilement maîtrisables. En effet un mauvais choix des critères d'homogénéité, qui ne sont définis qu'au début de la segmentation, peut conduire soit à une sous-segmentation soit à une sur-segmentation.

### 3.2 Un système immuno-multi-agents pour le traitement d'images biologiques

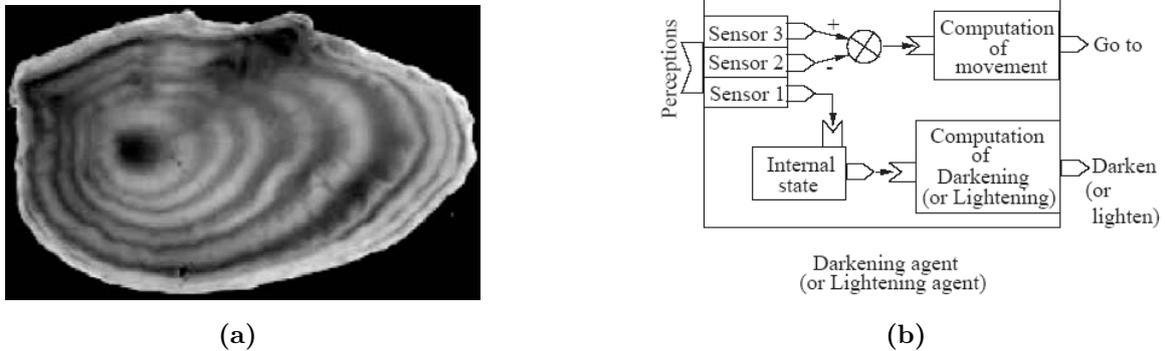
En utilisant le langage multi-agents oRis [HTRC02], Rodin et al. ont proposé un système multi-agents formé d'agents réactifs pour la détection de contours dans des images biologiques [RBG04]. En fonction de certains a priori sur le contenu des images, le système vise à produire une détection de contours meilleure que celle obtenue par les détecteurs traditionnels. Deux types d'agents dits respectivement agents de noircissement et agents de luminosité suivent respectivement les régions sombres et les régions claires. Leurs actions visent à renforcer les régions par l'accroissement de leur contraste, permettant ainsi une détection fiable de ces régions. Un immuno-mécanisme de régulation des systèmes multi-agents était utilisé pour détecter l'état final de la dynamique du système.

#### Architecture des agents

Pour une application de détection d'anneaux concentriques, tels que les anneaux d'estimation de l'âge de certaines espèces biologiques (Figure II.3a), les agents avaient une architecture réactive représentée à l'aide du formalisme BRIC - Block-like Representation of Interactive Components - [Fer99]. Chaque agent dispose de deux actuateurs et de trois capteurs (Figure II.3b).

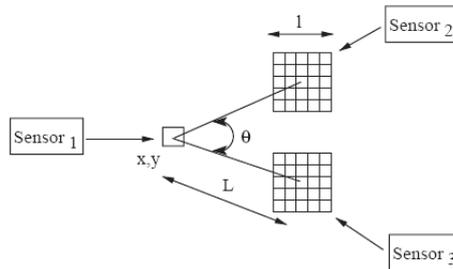
## II.3 Systèmes multi-agents pour l'analyse et l'interprétation d'images

Le premier actuateur permet à l'agent de faire des rotations autour de lui-même, ou d'avancer dans l'image. Le second lui permet d'agir sur l'image, en augmentant ou en diminuant l'illuminance d'un pixel, dans l'objectif de renforcer l'illuminance ou le noircissement de l'anneau associé. Les capteurs permettent à l'agent de mesurer l'illuminance des pixels situés dans son voisinage. Le premier (*sensor1*) lui permet de savoir s'il est situé sur un anneau déjà détecté ou non. Les deux autres (*sensor2* et *sensor3*), permettent à l'agent de mesurer les illuminances moyennes des deux zones de pixels situées à une



**Figure II.3** – Agents réactifs pour la détection d'anneaux concentriques. (a) Plaiice otolith de 8 ans ; (b) Architecture BRIC d'un agent

certainne distance ( $L$ ) en avant et écartées d'un certain angle ( $\theta$ ) (Figure II.4). L'agent, en calculant l'écart des deux moyennes, est capable de maintenir son parcours sur l'anneau qu'il lui y était associé.



**Figure II.4** – Capteurs d'un agent

### Dynamique du système

En fonction de l'information reçue de l'environnement à l'aide de ses capteurs, et de son état interne, l'agent prend une décision et l'exécute par le moyen de ses actuateurs. Son action consiste à renforcer l'anneau sur lequel il se déplace en augmentant l'illuminance des pixels correspondants lorsqu'il s'agit d'un anneau clair, ou en diminuant l'illuminance lorsqu'il s'agit d'un anneau sombre. L'agent s'arrête et s'auto-détruit lorsqu'il revient à sa position initiale, après avoir effectué le tour de l'anneau, en fournissant le polygone correspondant à son parcours.

### Discussion

Dans cette approche, les agents parcourant les différents contours sont totalement indépendants, et ne présentent aucune forme d'interaction ou de coopération entre eux. Le système paraît comme un algorithme parallèle de segmentation qui a été bien optimisé pour la détection des contours de type toit (roof edges) dans certains types d'images. Cependant, il peut échouer à détecter des contours plus simples tels que les contours de type marche. De plus, le nombre et la topologie des régions attendues doivent aussi être préalablement connus et figés dans les agents. Ceci est une limitation majeure de l'approche, car elle dépend fortement de la forme des objets qui peuvent exister dans les images traitées.

### 3.3 Surveillance multi-agents de scènes dynamiques

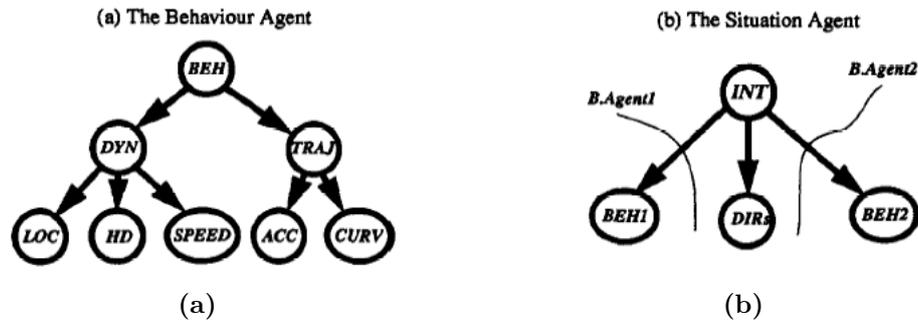
Les auteurs de ce travail [RTB98] présentent un système multi-agents pour la surveillance de scènes contenant des objets animés (piétons dans un parking) et des objets inanimés (véhicules stationnés). Pour chaque objet identifié de la scène, un agent est créé et y lui est affecté. Le système permet de suivre les objets animés et d'établir des associations en terme d'interaction entre les différents pairs d'objets situés en proximité les uns aux autres. Le système multi-agents, ainsi conçu, et en se basant sur l'information fournie par un système de détection, permet d'inférer l'interprétation la plus probable de la dynamique de la scène.

### Agents du système

Les auteurs utilisent un réseau bayésien afin de représenter la dynamique des objets suivis, et l'utilisent ensuite pour inférer la meilleure (au sens bayésien) interprétation de

## II.3 Systèmes multi-agents pour l'analyse et l'interprétation d'images

la dynamique des objets. Selon le comportement qui exhibent et le niveau dans lequel ils opèrent, les agents du système sont divisés en deux catégories :



**Figure II.5** – Réseau bayésien des agents utilisés : (a) Agent de comportement ; (b) Agent de situation.

### Agent de comportement :

Ce type d'agents opère au niveau objet. L'agent affecté à un objet donné de la scène est tenu de fournir une description de la cinématique de l'objet. Cette description est représentée par un réseau bayésien (Figure II.5a) et exprime en premier lieu la dynamique (DYN) de l'objet en terme de sa position (LOC) et de sa vitesse (HD : head speed). En second lieu, la trajectoire de l'objet (TRAJ) est exprimée par son accélération (ACC) et sa courbure (CURV). Les nœuds cachés du réseau DYN et TRAJ représentent une interprétation de l'objet associé en terme de dynamique (DYN) et de trajectoire (TRAJ).

### Agent de situation :

Ce type d'agents opère au niveau inter-objets. Pour chaque paire d'objets voisins une connexion probabiliste est créée et évaluée (Figure II.5b). Le nœud DIRS dans la Figure II.5b représente la direction des deux objets BEH1 et BEH2. Elle exprime le fait que les deux objets correspondants sont en direction l'un vers l'autre, l'un passant à proximité de l'autre, ou sont dans la même direction.

### Discussion

L'utilisation d'une approche bayésienne pour l'estimation des grandeurs exprimant la dynamique des objets dans la scène semble très pertinente du fait de l'ambiguïté d'in-

## II.3 Systèmes multi-agents pour l'analyse et l'interprétation d'images

---

interprétation des scènes dynamiques. Cependant, l'utilisation d'une approche multi-agents n'as pas été bien motivée. En effet, à part la distribution des traitements, en affectant à chaque objet un agent responsable de son suivi, les interactions entre les agents de comportement sont inexistantes. Pour les agents de situation, il n'a pas été explicité comment interagissent ces derniers entre eux, et où et comment l'information qui concerne l'interaction entre les objets est représentée et traitée.

### 3.4 Un système co-évolutionnaire pour la segmentation d'images

Le travail publié par Chen et Chen [CC05], fait partie d'un ensemble vaste de travaux ayant utilisé l'inspiration de la vie naturelle, notamment les phénomènes sociaux d'insectes, pour concevoir des systèmes robustes et parfois à fonctionnalité émergente pour l'analyse et la compréhension d'images [OB02]; [GKG06]; [CC05]. Les auteurs ont proposé un algorithme co-évolutionnaire pour la segmentation d'images. Les agents utilisés sont placés arbitrairement dans l'image, considérée comme un écosystème. Au cours de l'évolution, les agents s'organisent en tribus et échangent entre eux les propriétés locales de l'image, perçues dans leur voisinage. Le but de la co-évolution des agents est de minimiser l'intra-variabilité des propriétés locales au sein de chaque tribu et de maximiser l'inter-variabilité de ces propriétés entre les différentes tribus. A l'équilibre du système, les tribus représentent les différentes régions de l'image.

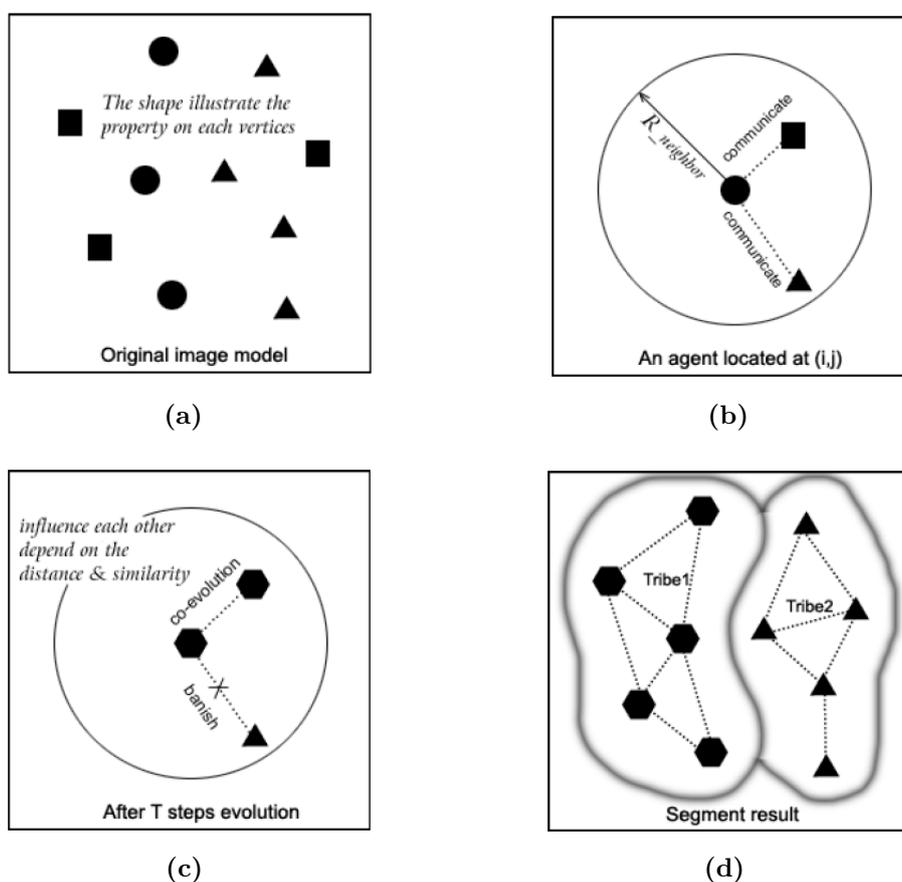
#### Modélisation de l'environnement

L'image à segmenter est considérée comme un écosystème, où les segments de l'image correspondent aux différentes tribus qui habitent l'environnement. Les agents représentent les individus de l'écosystème, où chacun est caractérisé par son identité, qui est représentée par un vecteur de propriétés. Les propriétés d'un agent sont celles du pixel sur lequel il est situé, à savoir : le niveau de gris, l'illuminance, la texture, etc.

Le long de l'évolution, chaque agent communique avec ses voisins afin d'échanger avec eux les informations d'identité (Figure II.6). En communiquant, chaque agent essaye d'influencer les autres, en ajustant en même temps son identité par une mise à jour de son vecteur de propriétés. Après plusieurs cycles (époques), les vecteurs de propriétés des agents tendent à devenir constants, et se forment des tribus stables. A la convergence, chaque tribu correspond à une région de l'image.

### Mécanisme d'évolution

Le processus d'évolution se déroule en plusieurs époques. A chaque époque, un agent qui est situé sur un pixel donné, communique avec les agents situés sur les pixels voisins, et échange l'information avec eux en utilisant des poids de communication proportionnels aux distances entre les agents. Afin de ne pas lisser les frontières entre les segments distincts (contours), deux agents ne s'influencent mutuellement que si la différence d'identité, calculée par l'écart entre les vecteurs de propriétés correspondants, est inférieure à un certain seuil (Figure II.6c).



**Figure II.6** – Co-évolution des tribus. (a) Identités initiales des agents ; (b) Echange d'identités entre agents ; (c) Influence mutuelle entre agents ; (d) Tribus stables et homogènes

### Discussion

Les systèmes inspirés de la vie naturelle sont bien connus pour leur robustesse et leur capacité d'auto-adaptation, aux imprévus qui peuvent surgir lors de l'interprétation des images. Ils sont également caractérisés par la propriété d'émergence, qui peut concerner des structures figurant dans l'image ou d'une fonctionnalité recherchée. Cependant, les résultats obtenus par ces systèmes sont généralement imprécis et ne peuvent pas être directement utilisés pour réaliser des mesures exactes ou pour fournir une interprétation fiable du contenu des images.

## 4 Conclusion

Dans ce chapitre, nous avons présenté une revue des travaux les plus répandus ayant utilisé l'approche multi-agents pour l'interprétation d'images et la reconnaissance d'objets. Nous avons essayé de diversifier l'analyse afin de montrer tous les aspects caractérisant ces systèmes. Les auteurs de ces approches ont proposé différentes stratégies pour une bonne utilisation des systèmes multi-agents afin de surmonter certains problèmes relatifs à l'interprétation d'images. Ils se sont différenciés également quant à l'architecture des agents : certains ont opté pour les agents réactifs, tandis que d'autres ont opté pour des agents cognitifs.

---

---

# Chapitre III

---

## Approches Bayésiennes en vision par ordinateur

### 1 Introduction

L'un des enjeux principaux dans le domaine de la recherche en Intelligence Artificielle est d'être capable de concevoir et de développer des systèmes dynamiques et évolutifs. De ce fait, ces derniers doivent être équipés de comportements intelligents qui peuvent apprendre et raisonner. Mais dans la plupart des cas, la connaissance acquise n'est pas toujours adéquate pour permettre au système de prendre la décision la plus appropriée. Pour pallier à ce problème, plusieurs méthodologies ont été proposées, mais seules les approches probabilistes s'adaptent mieux non seulement au raisonnement avec la connaissance et la croyance incertaine, mais aussi à la structure de la représentation de la connaissance. Ces approches probabilistes sont appelées "*réseaux bayésiens*" [Pea88]; [HM81]. Dans ce chapitre nous commençons par introduire les réseaux bayésiens et leurs applications

### 2 Réseaux Bayésiens

#### 2.1 Probabilité conditionnelle

Etant donné que l'événement  $B$  s'est produit, la probabilité que l'événement  $A$  se produise est  $X$  et s'écrit :  $P(A|B) = X$  .

## 2.2 Théorème de Bayes

Etant donnés deux événement  $A$  et  $B$  qui sont conditionnellement dépendants et un contexte  $c$ , le théorème de Bayes peut s'énoncer comme ci-dessous [SC94] :

$$P(B|A,c) = \frac{P(B|c)P(A|B,c)}{P(A|c)} \quad (\text{III.1})$$

où

- $P(B|A,c)$  est la probabilité *a posteriori* ou la probabilité de  $B$  après avoir pris en compte l'effet de  $A$  dans un contexte  $c$ ,
- $P(B|c)$  est la probabilité *a priori* de l'événement  $B$ , dans le contexte  $c$ ,
- $P(A|B,c)$  est la probabilité de  $A$  si l'on suppose que  $B$  est vrai dans un contexte  $c$ . Elle est appelée aussi la "vraisemblance",
- $P(A|c)$  est la normalisation, c'est la probabilité de l'événement  $A$  dans le contexte  $c$ .

## 2.3 Réseaux bayésiens

Un réseau bayésien est un graphe acyclique orienté sans circuit possible. Chaque nœud d'un réseau bayésien porte une étiquette qui est un des attributs du problème. Ces attributs sont binaires, pouvant prendre (avec certaine probabilité) la valeur *VRAI* ou *FAUX*, ce qui signifie qu'une variable aléatoire est associée à chaque attribut. Sur la figure III.1, on voit que  $A$  et  $C$  sont conditionnellement indépendants, on peut donc dire que  $P(A|B,C) = P(A|B)$ , la probabilité de  $A$  ne dépend que de  $B$ . Autrement dit, dans

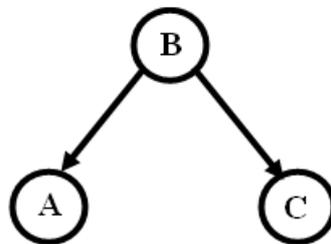


Figure III.1 – Un réseau bayésien

le réseau bayésien tout nœud est conditionnellement indépendant de ses nondescendants sachant ses parents [CM81]. La probabilité de  $C$  est donc  $P(C|B,A) = P(C|B)$ .

Dans cet exemple, la probabilité jointe de toutes les variables est :

$$P(A, B, C) = P(A|B) \times P(B) \times P(C|B)$$

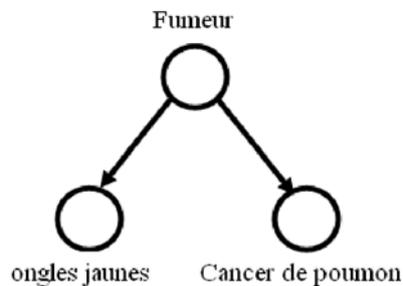
En général, étant donné les nœuds  $X_1, \dots, X_n$ , la probabilité jointe pour les réseaux bayésiens est :

$$P(X) = \prod_{i=1}^n P(X_i | \text{parents}(X_i))$$

La probabilité jointe de toutes les variables est le produit de toutes les probabilités de chaque variable et ainsi les valeurs de ses parents.

### 2.4 La notion de causalité

Dans les réseaux bayésiens, deux nœuds qui représentent deux faits différents peuvent être en relation causale sans que l'un implique l'autre. Graphiquement, chaque parent a un effet sur ses fils. La notion de causalité joue un rôle très important pour construire les réseaux bayésiens [CM81].



**Figure III.2** – Exemple de causalité

Par exemple sur un patient, si on constate qu'il a des ongles jaunes, alors on pourrait dire qu'il fume et en même temps, s'il fume, il a une forte chance d'avoir le cancer de poumon. Cela peut être représenté sous forme d'un graphe de causalité (Figure III.2).

### 2.5 Les inférences

Une fois qu'un réseau bayésien a été construit pour rendre compte d'un domaine, on cherche souvent à l'utiliser pour déterminer des probabilités correspondant à certains événements, certaines questions, certaines dépendances [CM81]. Dans ces cas, les inférences permettent effectivement de calculer la probabilité de chaque nœud lorsque les

autres variables sont inconnues [Ste03]. Il existe cependant quelques méthodes permettant d'effectuer ce calcul telles que *Conditionnement global*, *Arbre de jonction*, *Méthode approchée*.

## 2.6 Apprentissage

Lors de la construction de réseaux bayésiens, il n'est pas toujours évident qu'un expert puisse fournir de façon numérique l'ensemble des paramètres nécessaires à l'inférence dans un graphe. Il est intéressant dans certains cas de déterminer ces paramètres à partir d'une base d'exemples. C'est grâce à la méthode d'apprentissage qui permet de remédier ce problème. Les deux méthodes utilisées sont *apprentissage de paramètres* et *apprentissage de structure*.

## 3 Les applications

### 3.1 Domaines d'applications

Les réseaux bayésiens sont utilisés pour le diagnostic et le raisonnement dans des systèmes complexes dans de nombreux domaines :

- Marketing/Finance (modélisation de risques) :
  - Détection de fraudes
  - Optimisation de portefeuilles
- Informatique :
  - Différents assistants de résolution de problèmes
  - Evaluation de la qualité pour des développements logiciels
  - Reconnaissance de la parole, d'objets 3D
  - Diagnostic de pannes
- Médecine :
  - Aide au diagnostic médicale
  - Surveillance transfusionnelle
- Industrie :
  - Aide au diagnostic de pannes en temps réel
  - Système de contrôle d'un véhicule
  - Aide au télédiagnostic
  - Modélisation de groupes électrogènes

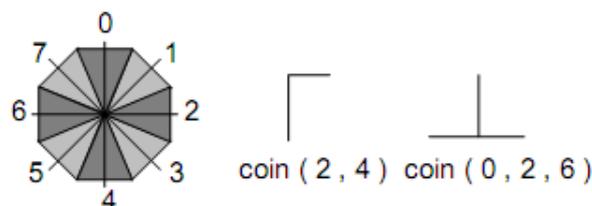
- Environnement :
  - Prévisions météo
  - Prédiction des pics d’ozone

## 4 Réseaux Bayésiens en vision par ordinateur

### 4.1 Reconnaître des modèles géométriques simples

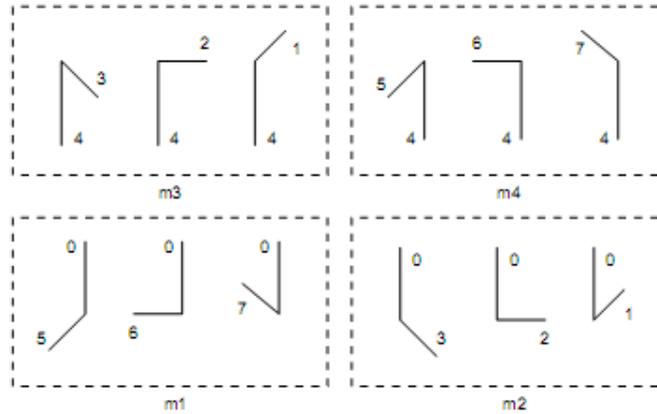
Djian, Probert et Rives utilisent les réseaux bayésiens dans un cadre de reconnaissance d’objets [DRP96]. Ils cherchent à reconnaître des modèles géométriques simples (des portes et des fenêtres) dans une séquence d’images pour cela, des observateurs dynamiques cherchent à détecter l’apparition de motifs élémentaires dans la scène. Les réseaux bayésiens servent ici à guider le raisonnement en prenant une décision sur l’action à exécuter (positionnement dans l’image des observateurs) et en émettant une hypothèse sur le motif que l’on doit détecter dans ces observateurs. Un deuxième réseau permet de vérifier si l’hypothèse émise est correcte.

Dans l’image un coin est défini comme la jonction de deux segments de droites ou plus (intersection des contours d’une classe d’objet). Pour des raisons de simplicité, les auteurs ont discrétisés les orientations possibles d’un segment selon le schéma de la figure III.3.



**Figure III.3** – Discretisation de l’orientation des segments et exemple de coins.

A partir des primitives de type coin, on peut regrouper les différents types de coins observables dans les occurrences de chacune des deux classes en quatre motifs élémentaires que nous appellerons  $m_1$ ,  $m_2$ ,  $m_3$ , et  $m_4$ . Ceux-ci sont représentés sur le schéma de la figure III.4 qui prend en compte les différences de perspectives associées aux différents angles de vue de la caméra par rapport à l’objet. Ainsi le motif  $m_1$  pourra représenter le coin inférieur gauche d’une porte ou le coin inférieur droit d’une fenêtre.



**Figure III.4** – Motifs élémentaires : coins à jonction double.

Pour traiter les problèmes des incertitudes et les erreurs de modèles, trois réseaux bayésiens ont été proposées :

- Génération d’hypothèse (réseaux hypothèse) ;
- Vérification d’hypothèse (réseau vérification) ;
- Instanciation des modèles (réseaux modèles).

#### Réseau hypothèse :

Ce réseau sur la figure III.5 permet d’effectuer le raisonnement suivant : (i) identification du motif courant, (ii) choix de la prochaine action, et (iii) prédiction du motif suivant :

- Identification : l’instanciation des nœuds correspondants aux segments détectés dans une région d’intérêt de l’image permet d’identifier le motif courant ;
- Prise de décision : le motif courant et notre croyance sur l’objet présent dans l’environnement permettent de déterminer un ensemble d’actions intéressantes ; si nous pensons voir une porte, et que nous venons de détecter le motif  $m_1$ , alors :
  - La meilleure action consiste à regarder vers le *Nord* pour suivre le segment  $S_0$  de  $m_1$  ;
  - Regarder vers l’*Est* est acceptable bien que moins désirable car suivre l’opposé du segment  $S_6$  de  $m_1$  sera plus risqué (pas de segment sur lequel s’appuyer) ;
  - Regarder vers l’*Ouest* ou vers le Sud n’apportera rien à la reconnaissance d’une porte.

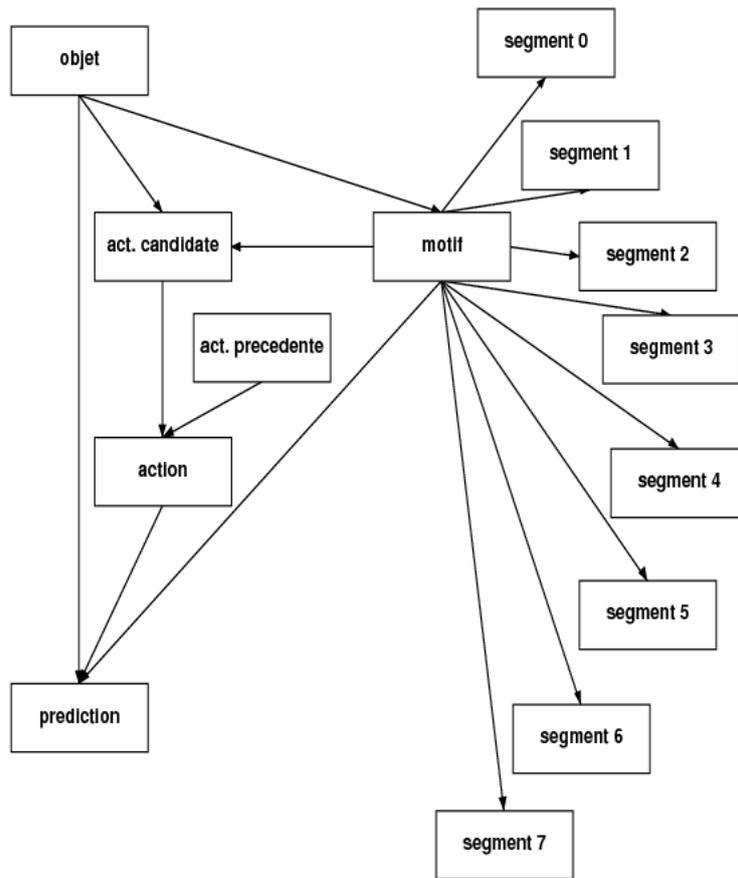


Figure III.5 – Réseau hypothèse

Nous pouvons traduire la phrase précédente en terme de probabilités conditionnelles :

$$P(\text{candidat} = N / \text{objet} = \text{porte}, \text{motif} = m1) = 0.55$$

$$P(\text{candidat} = S / \text{objet} = \text{porte}, \text{motif} = m1) = 0.05$$

$$P(\text{candidat} = E / \text{objet} = \text{porte}, \text{motif} = m1) = 0.35$$

$$P(\text{candidat} = O / \text{objet} = \text{porte}, \text{motif} = m1) = 0.05$$

- Prédiction : la troisième étape consiste à prédire le motif suivant une fois que la meilleure action a été choisie

#### Réseau de vérification :

Ce réseau de la figure III.6 permet de savoir si aucune, l'une des deux seulement, ou si les deux hypothèses ont été vérifiées à partir des mesures courantes et les segments détectés dans une région d'intérêt de l'image.

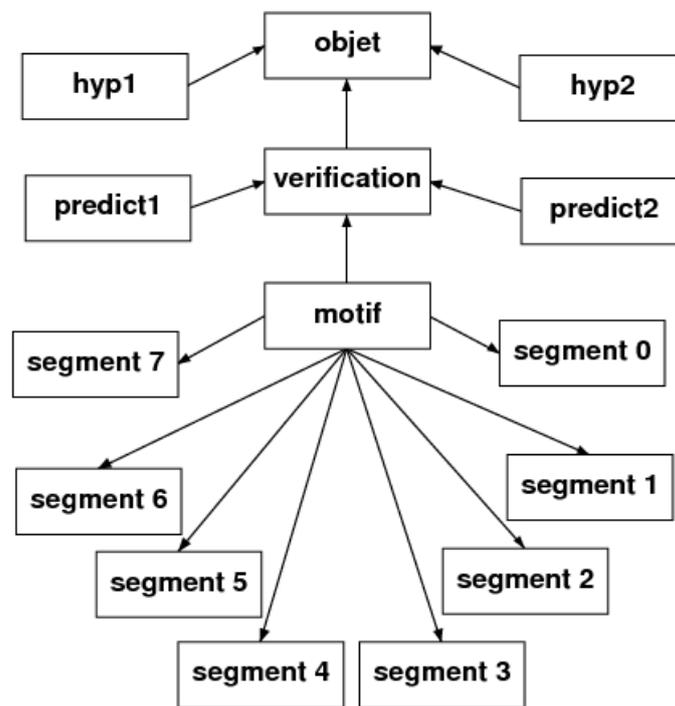


Figure III.6 – Vérification d'hypothèse

#### Réseau de modélisation :

Ces réseaux servent à maintenir la connaissance sur les motifs effectivement vus à l'endroit attendu pour chacun des modèles et à détecter la fin de la tâche de reconnaissance lorsque tous les motifs d'un objet ont été vus là où ils étaient attendus.

Le réseau de la figure III.7 permet de maintenir la connaissance sur la présence éventuelle d'une fenêtre dans l'environnement.

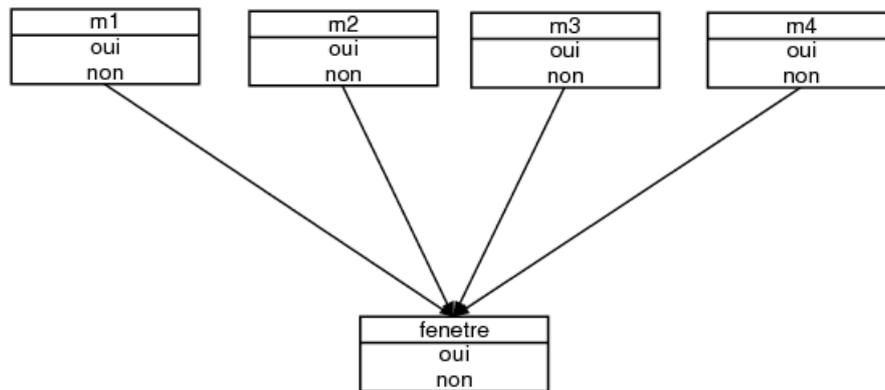


Figure III.7 – Modèle d'une fenêtre

#### Discussion

Ce système de reconnaissance est efficace sur des modèles géométriques simples. Les auteurs peuvent distribuer le processus des raisonnements (prise de décision) selon une approche de type multi-agents dans laquelle les observateurs dynamiques, les modèles et le contrôleur coopèrent à la reconnaissance.

#### 4.2 Le contrôle de la perception sélective

Rimey et Brown [RB94] ont développé le concept de vision sélective basée sur la collecte séquentielle d'information pour répondre à des questions spécifiques. Ces questions concernent l'interprétation d'une scène de manière générale ou bien encore la détection de la présence d'un objet donné. Le système proposé n'exécute que les actions nécessaires pour pouvoir répondre à la question posée, et utilise une base de connaissances sur le modèle général de la scène (organisation) et sur les modèles particuliers des objets. L'ensemble de ces connaissances est regroupé dans des réseaux bayésiens. Certains de ces réseaux regroupent la connaissance sur un domaine spécifique (*composite nets*) d'autres regroupent la connaissance relative à une tâche précise (*task net*). Le composite net est lui-même décomposé en trois réseaux bayésiens, l'un d'eux regroupe la structure physique de la scène (*PART OF net*) indiquant la présence ou l'absence de tel ou tel objet dans la scène, un autre regroupe les relations géométriques (*expected area net*), c'est-à-dire où l'objet devrait se situer dans le plan observé, le dernier, enfin, exprime les relations hiérarchiques entre objets ou groupes d'objets (réseau en forme d'arbre *IS-A tree*). Les actions sont séparées en deux catégories en fonction de leur coût : d'une part les actions visuelles

traitant les données extraites des images avec pour objectif l'extraction d'indices visuels nécessaires pour répondre aux questions posées, d'autre part les actions de déplacement de la caméra (calcul de points de vue) dont le coût est plus élevée. Le système décrit utilise les informations contenues dans l'ensemble des réseaux bayésiens et les fonctions de coût pour déterminer la meilleure séquence d'actions à exécuter. Deux solutions de planification des tâches sont comparées, la première est basée sur le calcul d'une fonction de qualité qui décide de la nécessité de déplacer la caméra ou d'effectuer une action visuelle. La seconde consiste à énumérer l'ensemble des séquences d'actions réalisables et d'extraire celle dont le coût est le plus faible.

#### Discussion

La méthode employée ci-dessus reste dédiée à des tâches très spécifiques qui nécessitent de définir précisément les *a priori* contenus dans les différents réseaux bayésiens.

### 4.3 Surveillance du trafic routier

Buxton et Gong, [BG95] cherchent à construire un système de surveillance du trafic routier. Ils ne souhaitent pas uniquement suivre les véhicules dans une séquence d'images mais ils désirent aussi analyser, interpréter et classier leur comportement. Pour cela, il convient de disposer de connaissances *a priori* sur les comportements possibles des différents objets mis en jeu. De plus, il est nécessaire de concevoir un système permettant de définir les différentes actions de perception utiles à l'accomplissement de la tâche de surveillance. Les réseaux bayésiens fournissent un cadre unifié permettant de disposer à la fois de la connaissance *a priori* et de la modélisation du contrôle de tâche. Enfin, ils autorisent une gestion dynamique (grâce au processus de propagation) des dépendances entre les acteurs impliqués dans le processus d'interprétation visuelle. L'approche proposée est décomposée en trois points fournir une interprétation possible (la plus probable) des données disponibles puis mettre à jour cette interprétation en fonction des nouvelles données et de l'interprétation courante.

#### Discussion

Dans cet article, Buxton et Gong proposent un cadre méthodologique général intéressant pour la conception et la mise en œuvre de systèmes de surveillance du trafic routier.

## 5 Conclusion

Le théorème de Bayes énoncé au 18<sup>ème</sup> siècle s'est révélé être un outil très puissant de décision ou d'aide à la décision en informatique. Par une approche statistique et probabiliste des données, les réseaux bayésiens permettent d'établir un formalisme des connaissances incertaines, ce qui les place au cœur du domaine l'intelligence artificielle.

Les travaux que nous avons présentés, illustrent bien les orientations des recherches d'aujourd'hui sur les réseaux bayésiens, telle que leur utilisation en reconnaissance des modèles géométriques, contrôle routier et perception sélective.

Dans le chapitre suivant, nous abordons notre contribution à travers ce mémoire. Il s'agit de la conception d'un système de reconstruction et de reconnaissance d'objets basé à la fois sur un système multi-agent et sur l'approche d'inférence bayésienne.

---

---

# Chapitre IV

---

## Une approche bayésienne distribuée pour la reconnaissance d'objets

### 1 Introduction

Nous exposons dans ce chapitre une approche bayésienne à base d'agents pour la reconnaissance d'objets polyédriques. En premier temps une population d'agents détecteurs lancée au premier niveau du système pour l'extraction des segments de droites, ces derniers sont envoyés au deuxième niveau du système pour les reconstruire par l'agent bayésien, et en fin les agents reconnaisseurs qui opèrent dans le troisième niveau et prennent en charge la tâche de reconnaissance d'objets.

### 2 Description de l'approche

Nous proposons dans ce chapitre une approche de reconnaissance d'objets. Cette dernière utilise un système multi-agents cognitif pour la détection/reconstruction de primitives issues de la segmentation collective d'images de profondeur proposée par Mazouzi et al. [MGM09]. Dans le premier niveau de notre système décrit dans la figure IV.1, une population d'agents dits *agents détecteurs* évolue sur l'image et suit les points de contour dont le but est de former des segments de droites, et les enregistrer dans une structure de données nommée  $T$ . Un autre agent nommé *agent bayésien* opère dans le deuxième niveau de notre système, et qui exploite les connaissances *a priori* sur les scènes polyédriques (ces connaissances sont codées sur des réseaux bayésiens), afin de former des objets complexes à partir de segments détectés. Et enfin, un *agent reconnaisseur* qui prend en charge la tâche de reconnaissance et engendre le processus d'identification de type Objet-Modèle [HM95].

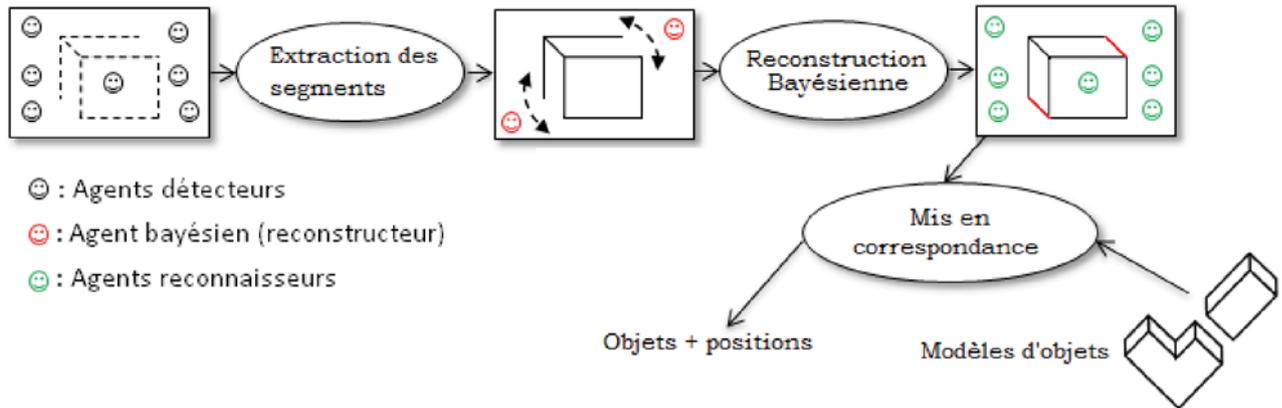


Figure IV.1 – Schéma de l'approche proposée

## 2.1 Agents détecteurs (Agents Père, Agents Fils)

Cette population d'agents dite *agents\_père*, travaille dans le premier niveau de notre système et prend la tâche de suivre les points de contour. Chaque agent de cette population est caractérisé par un ensemble d'attributs dont : Sa position courante  $(x,y)$  dans l'image, le rayon de détection  $R$  et les points des extrémités  $(x_d,y_d)$ ,  $(x_g,y_g)$ . Loin de contour un agent père se déplace aléatoirement (change aléatoirement sa direction). A la rencontre du premier point de contour, l'agent reproduit deux agents dits *agent\_fils* ( $A_d$  : Agent fils droit,  $A_g$  : Agent fils gauche), l'un pour suivre la partie droite du segment, et l'autre pour la partie gauche. Dès que les deux agents  $A_d$ ,  $A_g$  trouvent les points d'extrémités droite et gauche du segment, ces derniers envoient leurs positions à l'agent père. Dans le cas où l'agent père est placé sur un point isolé ou un point de coin, ce dernier s'autodétruit par lui-même (Figure IV.2).

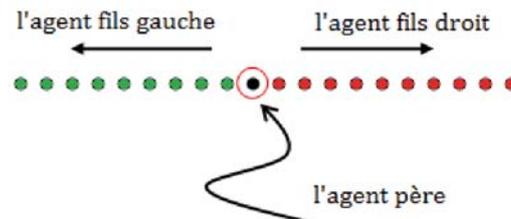


Figure IV.2 – Les positions des agents sur les points de contour

### Points isolés

Pendant le parcours des agents pères sur l'ensemble des points de contour, ils testent la présence des voisins de chacun de ces points, afin d'éliminer ceux qui ne participent pas à la construction des segments, c'est-à-dire les points *isolés*. pour accomplir cette tâche, nous utilisons un cercle de rayon  $R$  (Figure IV.3b) ( $R$  représente une distance entre un pixel central et ses voisins) pour calculer les coordonnées de voisinage d'un pixel central et un compteur *compt*, pour compter le nombre de ce dernier. Alors ces coordonnées sont données par :

$$\begin{aligned}x &= x + R * \text{Cos}(\theta) \\ y &= y + R * \text{Sin}(\theta)\end{aligned}$$

La fonction ci-dessous teste si un pixel est isolé ou non :

---

**Algorithme 2.1:** POINT\_ISOLÉ( $x : \text{entier}, y : \text{entier}$ )

---

**Commentaire:** Tab\_Pixels : le tableau des pixels de l'image

**variables entières** *compt, x<sub>1</sub>, y<sub>1</sub>*

**variables réelles**  $\theta$

*compt*  $\leftarrow 0$

$\theta \leftarrow 0$

**répéter**

$$\left\{ \begin{array}{l} x_1 \leftarrow x + R * \text{Cos}(\theta) \\ y_1 \leftarrow y + R * \text{Sin}(\theta) \\ \text{si } \text{Tab\_Pixels}(x, y) = 0 \\ \quad \text{alors } \text{compt} \leftarrow \text{compt} + 1 \\ \theta \leftarrow \theta + (2 * PI / 360) \end{array} \right.$$

**jusqu'à**  $\theta > 2 * PI$

**si** *compt* = 0

*Point\_Isolé*  $\leftarrow VRAI$

**sinon**

*Point\_Isolé*  $\leftarrow FAUX$

---

Points du coins

A chaque point de contour, les agents fils calculent l'angle  $\theta$  entre leurs positions courantes  $(x_d, y_d)$ ,  $(x_g, y_g)$  et la position de leur père  $(x, y)$ , si cette dernière est plus grande qu'un certain *seuil* les agents fils envoient leurs coordonnées (qui représentent les extrémités gauche et droite du segment) à l'agent père, afin de les enregistrer dans un tableau nommé  $T$ . dès qu'un agent fils envoi ses coordonnées à son père il termine sa tâche par l'autodestruction (Figure IV.3a).

On calcule l'angle entre deux points  $p_1(x_1, y_1)$  et  $p_2(x_2, y_2)$  par la formule suivante :

$$\theta = \arctan((y_2 - y_1)/(x_2 - x_1))$$

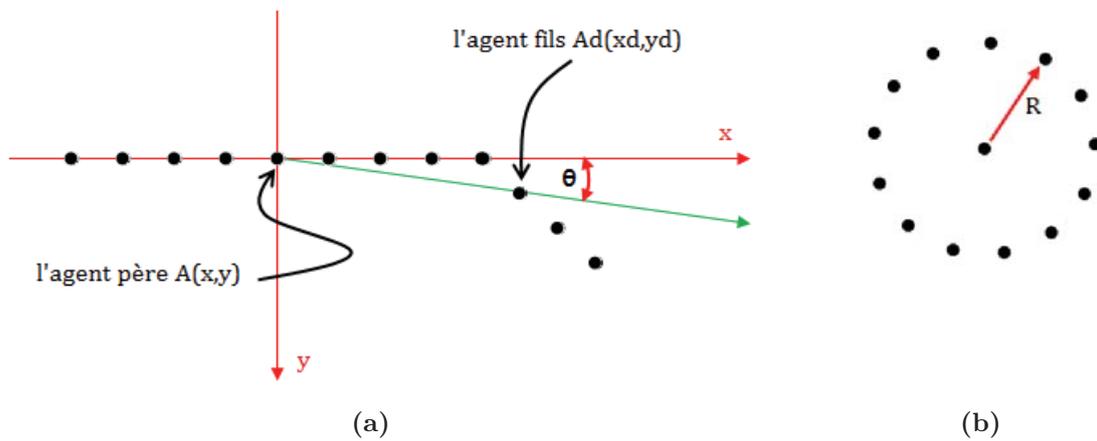


Figure IV.3 – Les deux types de points :(a) Point de coin ; (b) Point isolé.

Chaque agent père a le comportement suivant :

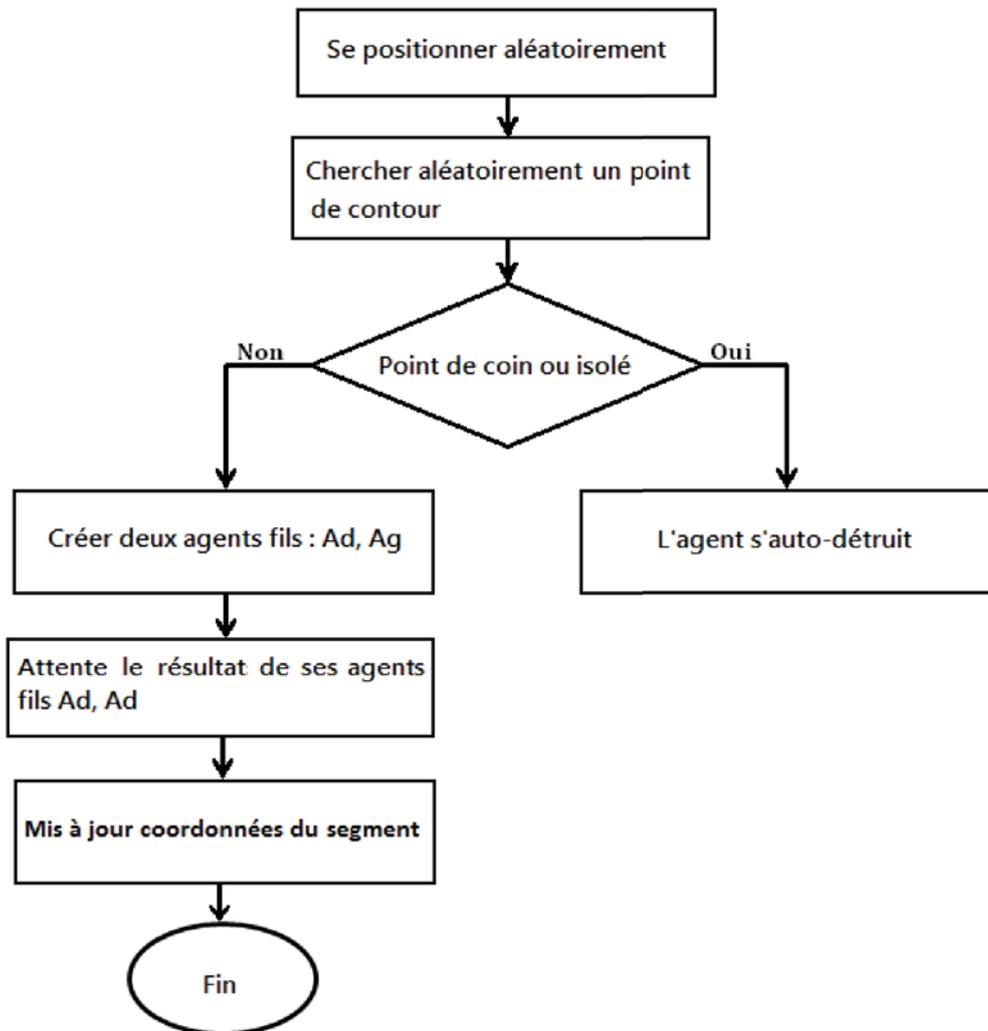


Figure IV.4 – Comportement de l'agent père

Chaque agent fils a le comportement suivant :

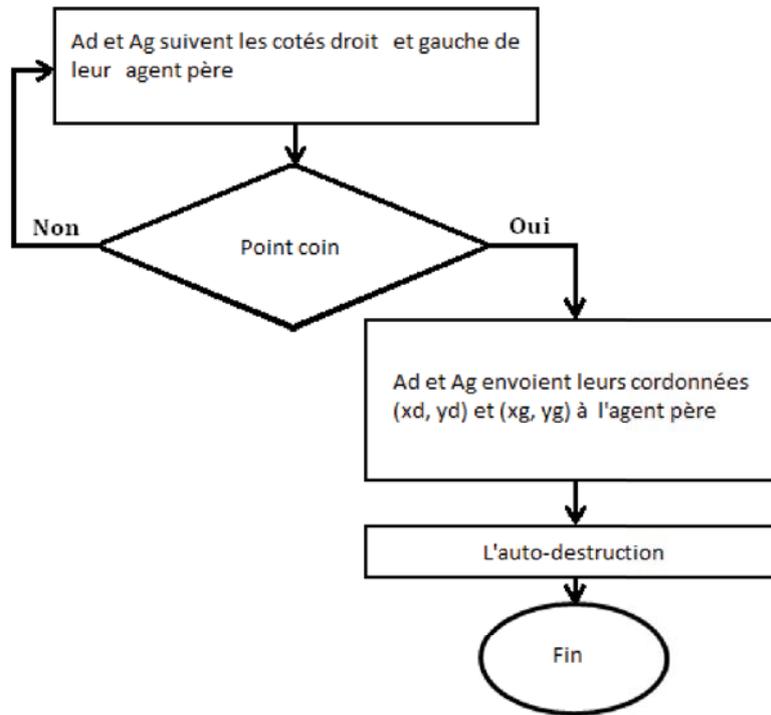


Figure IV.5 – Comportement de l'agent fils

Le pseudo-code ci-dessous représente la description algorithmique de la structure des agents et leurs méthodes. :

**Variables :**

TabAgents(NBagent) : Table de 1 à 100 d'Agent\_ père

NBagent : entier

Tab\_pixels(.) : Le tableau de l'image

La structure de l'agent père :

<b>Agent_père( )</b>
<b>Début</b>
<u>Propriétés</u>
$x, y, x_d, y_d, x_g, y_g, R$ : entier
<u>Méthodes</u>
Parcourir_Image( )
Lancer_Agents_Fils( )
Autodestruction( )
<b>Fin</b>

Cette fonction initialise aléatoirement les agents pères sur l'image

---

### Algorithme 2.2: INITIALISATION()

---

**Commentaire:** Initialiser les agents pères sur l'image

$N_{Agent} \leftarrow 100$

**pour**  $i = 0$  **jusqu'à**  $N_{Agent}$

**faire**  $\left\{ \begin{array}{l} Agentpr \leftarrow newAgent\_d\acute{e}tecteur() \\ Agentpr.x \leftarrow Random(image.width) \\ Agentpr.y \leftarrow Random(image.height) \\ TabAgents(i) \leftarrow Agentpr \end{array} \right.$

---

## IV.2 Description de l'approche

---

La méthode suivante permet aux agents pères de se déplacer aléatoirement sur l'image à la recherche des points de contour.

---

### Algorithme 2.3: PARCOURIR\_IMAGE()

---

**Commentaire:** La méthode parcourir image

**variables booléennes** *Arrêté*

**répéter**

$$\begin{cases} x_1 \leftarrow x + (1 - \text{Random}(2)) \\ y_1 \leftarrow y + (1 - \text{Random}(2)) \\ \text{Arrêté} \leftarrow \text{Tab\_pixels}(x, y) = 0 \end{cases}$$

**jusqu'à** *Arrêté = VRAIS*

*Agent\_détecteur.x*  $\leftarrow$  *x*

*Agent\_détecteur.y*  $\leftarrow$  *y*

---

La structure de l'agent Fils :

<b>Agent_Fils( )</b>
<b>Début</b>
<u>Propriétés</u>
<i>x<sub>e</sub>, y<sub>e</sub></i> : entier <i>/* les points d'extrémités</i>
$\Theta$ : Réel <i>/* L'angle entre l'agent père et l'agent fils</i>
<i>R</i> : entier <i>/* le rayon du cercle de voisinage</i>
<u>Méthodes</u>
Parcourir_Segment( )
Point_Coin( )
Autodestruction( )
<b>Fin</b>

Ci-après, la description algorithmique de la méthode qui permet aux agents fils de parcourir les deux parties gauche et droite du segment.

---

**Algorithme 2.4:** PARCOURIR\_SEGMENT( $x : ent, y : ent, cx : ent, cy : ent$ )

---

**Commentaire:**  $cx, cy$  : représentent les coordonnées de l'agent père

**variables entières**  $xp, yp, R$

**variables réelles**  $\theta$

**répéter**

$$\left\{ \begin{array}{l} \theta \leftarrow 0 \\ \text{répéter} \\ \left\{ \begin{array}{l} xp \leftarrow x + R * \text{Cos}(\theta) \\ yp \leftarrow y + R * \text{Sin}(\theta) \\ \text{si } \text{Tab\_Pixels}(xp, yp) = 0 \text{ et } \text{distance}(xp, yp, cx, cy) > \text{distance}(x, y, cx, cy) \\ \text{alors } \left\{ \begin{array}{l} x \leftarrow xp \\ y \leftarrow yp \end{array} \right. \\ \theta \leftarrow \theta + (2 * PI / 360) \end{array} \right. \\ \text{jusqu'à } \theta > 2 * PI \\ \text{Agent\_Fils.xe} \leftarrow x \\ \text{Agent\_Fils.ye} \leftarrow y \end{array} \right.$$

**jusqu'à**  $\text{coin}(x, y)$

---

## 2.2 Agent bayésien

Afin d'améliorer les résultats de reconstruction, obtenus par le système multi-agent décrit précédemment, nous nous sommes inspirés du travail d'ERIC Marchand [MC98] pour mettre en œuvre un agent bayésien pour la correction des résultats obtenus par les agents détecteurs. Cet agent a une seule instance et son objectif principal est d'obtenir une représentations de la scène en terme d'objets simple : segments, polygones. L'agent bayésien proposé pour parvenir à cet objectif repose sur des techniques de prédiction/vérification d'hypothèses. Pour cela, l'agent bayésien procède en deux étapes qui sont modélisées dans des réseaux bayésiens (création et vérification d'hypothèses). Dans un premier temps ce dernier parcourt la liste  $T$  des segments détectés, et pour chaque couple de segments  $(S_{t'}, S_t)$ , il détermine les informations 3D supplémentaires que ce couple de segments peut apporter. Puis vérifier que les hypothèses émises sont conformes à la réalité.

### 2.2.1 Prédiction : génération des hypothèses

Dans cette étape l'agent bayésien exploite un réseau bayésien (Figure IV.6) pour émettre une ou plusieurs hypothèses (concernant les relations entre les deux extrémités les plus proches des segments et l'autre concernant les relations entre les deux extrémités les plus éloignées) menant à la réalisation d'une ou plusieurs de ces actions suivante :

- Fusionner les deux segments  $S_{i'}$  et  $S_i$  ;
- Joindre les deux segments  $S_{i'}$  et  $S_i$  en une jonction  $J_{i'i}$  ;
- Ajouter un ou deux segments entre  $S_{i'}$  et  $S_i$ .

Pour cela, l'agent procède en plusieurs étapes : Dans un premier temps, il cherche des relations topologiques simples (*voisinage, coplanarité, colinéarité*) entre les deux segments  $S_{i'}$  et  $S_i$ , puis il définit cinq classes distinctes auxquelles peuvent appartenir ces segments, chacune de ces classes représentant une combinaison particulière des différentes relations élémentaires précédemment définies. Enfin, à partir de ces classes, notre agent peut émettre une hypothèse sur la nature des relations associées aux segments. Ce schéma de raisonnement est codé dans la structure du réseau bayésien de la figure IV.6. Celui-ci comporte six nœuds principaux, chacun étant associé à une étape du raisonnement.

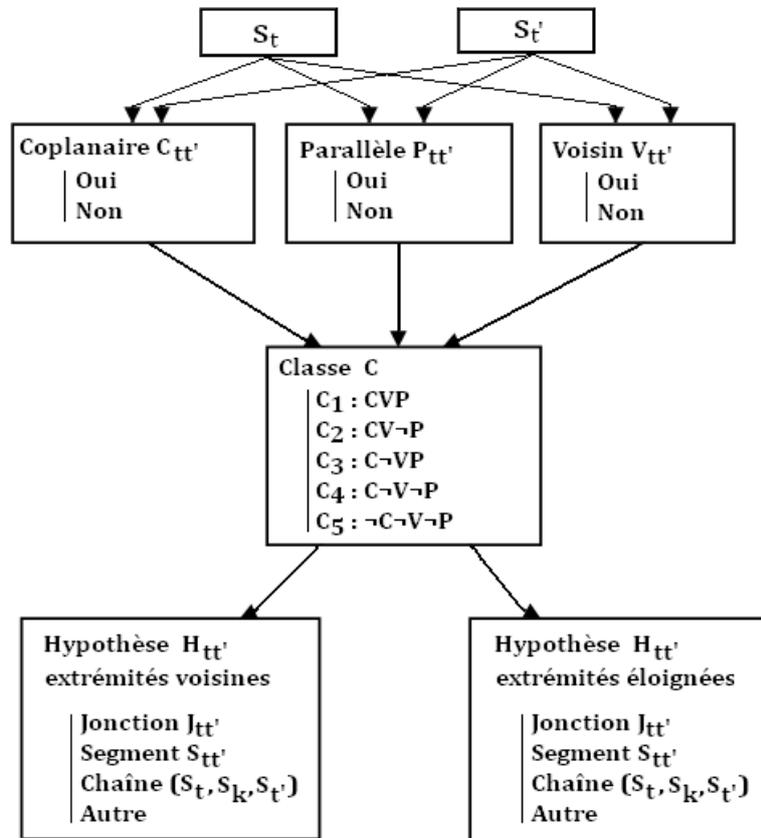


Figure IV.6 – Réseau de génération d'hypothèses

On peut classer les segments comme suivant :

- $C_1$  : les segments sont coplanaires, voisins et parallèles  $CVP$  ;
- $C_2$  : les segments sont coplanaires, voisins et non parallèles  $CV\neg P$  ;
- $C_3$  : les segments sont coplanaires, non voisins et parallèles  $C\neg VP$  ;
- $C_4$  : les segments sont coplanaires non voisins et non parallèles  $C\neg V\neg P$  ;
- $C_5$  : les segments sont non coplanaires, non voisins et non parallèles  $\neg C\neg V\neg P$ .

Enfin, à partir de la classe à laquelle appartient ce couple de segments, il convient de déterminer l'hypothèse à adopter concernant la nature des relations entre ces deux segments. Nous allons donc définir des stratégies de décisions qui détermineront la meilleure hypothèse possible, compte tenu des connaissances disponibles. Ces stratégies sont codées dans les tables de probabilités conditionnelles  $P(H|C)$ . Deux séries d'hypothèses sont émises en même temps, l'une concerne les relations établies entre les extrémités les plus proches des segments  $S_t$  et  $S_{t'}$  (Figure IV.7a) et l'autre les relations établies entre les

## IV.2 Description de l'approche

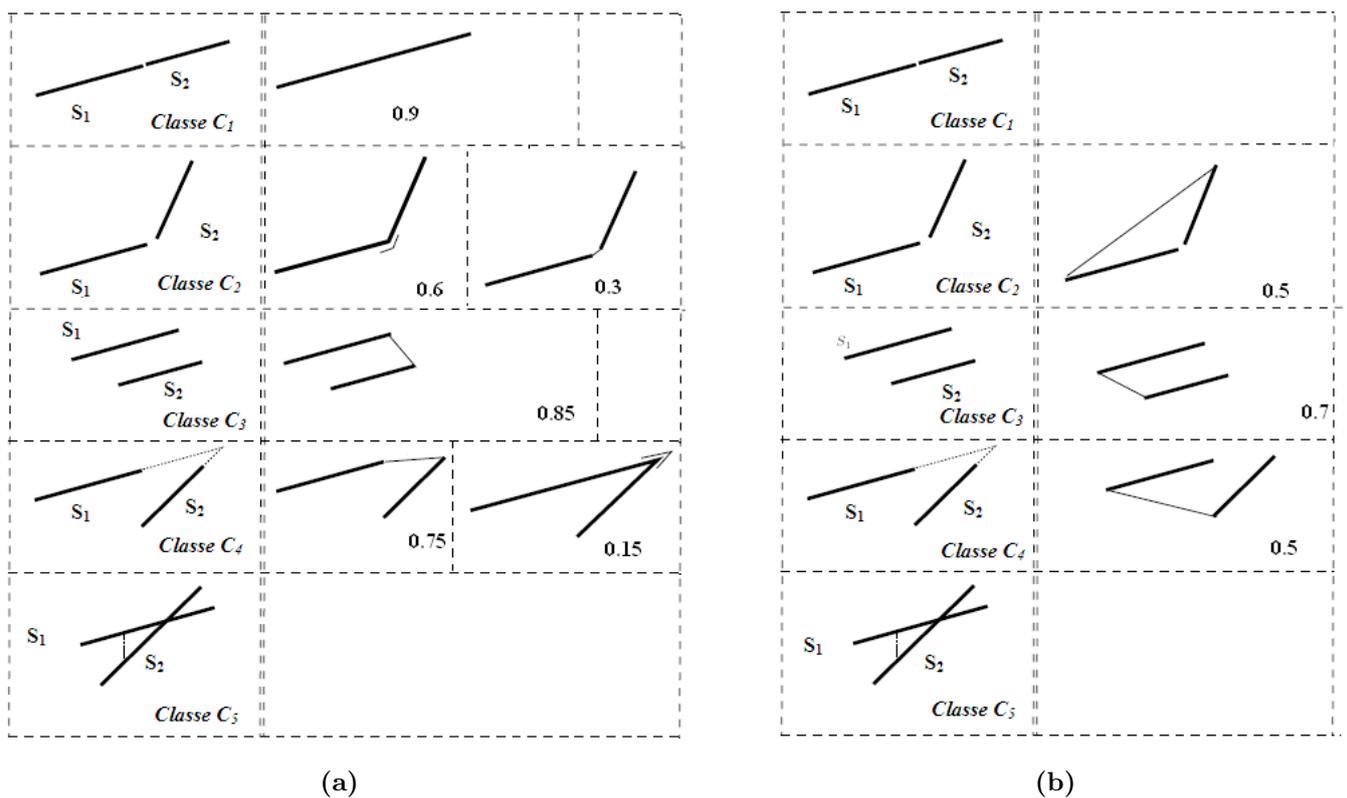
extrémités les plus éloignées (Figure IV.7b) . Si nous prenons l'exemple de deux segments voisins, coplanaires et non parallèles (ces deux segments appartiennent donc à la classe  $C_2$ , et les hypothèses associées sont décrites sur la deuxième ligne de la figure IV.7a . L'hypothèse la plus vraisemblable dans ce cas est de supposer la présence d'une jonction entre ces deux segments. On a donc la table de probabilités associées à cette classe :

$$P(H = fusion|C_2) = 0.025$$

$$P(H = jonction|C_2) = 0.6$$

$$P(H = chaîne|C_2) = 0.3$$

$$P(H = autre|C_2) = 0.075$$



**Figure IV.7** – Classe élémentaires et hypothèses associées :(a) Extrémités proches ;  
(b) Extrémités éloignés.

Les tableaux et les différentes probabilités conditionnelles associées aux hypothèses, sont présenté ci-après.

## IV.2 Description de l'approche

	Classe				
hypothèse	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
$H_1$ Segment	0.90	0.025	0.05	0.05	0.025
$H_2$ Jonction	0.025	0.60	0.05	0.15	0.025
$H_3$ Chaîne	0.025	0.30	0.85	0.75	0.025
$H_4$ Autres	0.05	0.075	0.05	0.05	0.925

**Tableau IV.1** – Table de probabilités conditionnelles  $p(\text{hypothèse}|\text{classe})$  pour les extrémités proches

	Classe				
hypothèse	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
$H_1$ Segment	0.025	0.025	0.025	0.025	0.025
$H_2$ Jonction	0.025	0.025	0.025	0.025	0.025
$H_3$ Chaîne	0.025	0.5	0.7	0.5	0.025
$H_4$ Autres	0.925	0.45	0.25	0.45	0.925

**Tableau IV.2** – Table de probabilités conditionnelles  $p(\text{hypothèse}|\text{classe})$  pour les extrémités éloignées

### 2.2.2 Vérification des hypothèses émises

L'étape de prédiction repose essentiellement, comme nous l'avons vu, sur les connaissances 3D déjà acquises et sur la connaissance d'un « expert » qui est codée dans un réseau bayésien. Celui-ci, présenté sur la figure IV.8 permet d'inférer la nature de l'objet associé aux deux segments considérés en confirmant ou infirmant les deux hypothèses ( $H_{Max}^1$  et  $H_{Max}^2$ ) favorisées dans l'étape de prédiction.

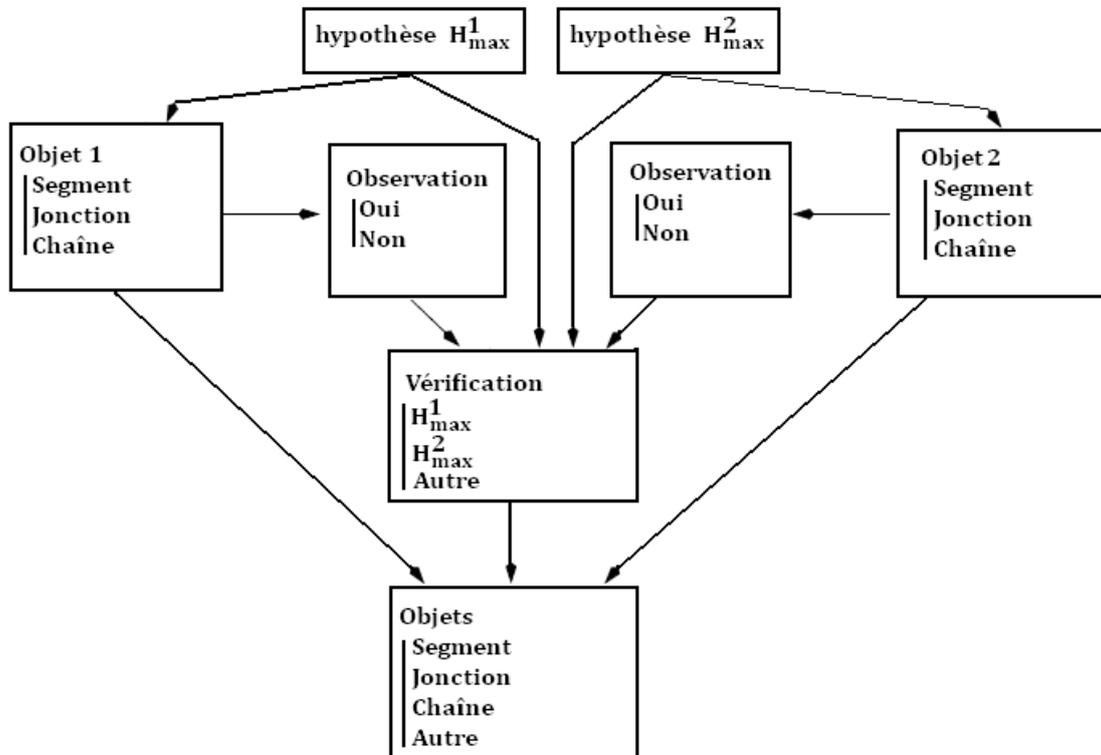


Figure IV.8 – Réseau de vérification d'hypothèses

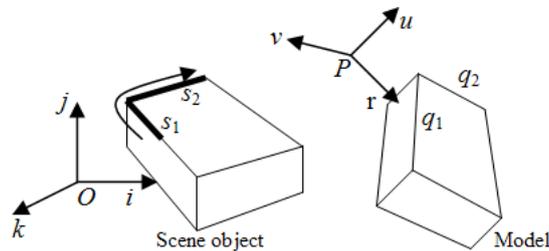
Dans un premier temps, en considérant les deux hypothèses retenues, nous définissons, la nature (*segment*, *jonction*, *chaîne*) et la position de l'objet qui sont associées à cette hypothèse. Puis, en utilisant des informations extraites des images, il est possible de déterminer si cet objet existe ou non. C'est l'étape la plus importante, car elle fait directement intervenir les stratégies de perception. En effet, si dans certains cas l'hypothèse peut être directement confirmée ou infirmée en se basant uniquement sur les informations déjà acquises, dans d'autres cas, cette information peut s'avérer insuffisante. Il est alors nécessaire de déplacer la caméra afin d'acquérir (ou non) le complément d'information attendu. Ainsi, connaissant la confiance accordée à chacune des deux hypothèses, ainsi que la vraisemblance des observations associées, il est possible de déterminer laquelle des deux hypothèses, était valide (ou, éventuellement de rejeter les deux hypothèses). Enfin, compte tenu de la validation de l'une de ces deux hypothèses, nous en déduisons l'objet associé au couple de segments considéré.

## 2.3 Les agents reconnaisseurs

Dans ce niveau deux types d'agents sont proposés, Le premier représente *les agents reconnaisseurs*, qui prennent en charge la tâche du suivi des contours, et l'établissement des hypothèses de correspondance. Le second type d'agent existe en une seule instance et représente l'*agent coordonnateur* qui examine et valide les différentes hypothèses de mise en correspondance, générées et envoyées par les agents reconnaisseurs.

### 2.3.1 Les agents reconnaisseurs :

Ce type d'agent a pour objectif de suivre les points du contours, qui forme les objets modélisés par l'agent bayésien, afin de générer les hypothèses de correspondances objets-Modèle. En dehors des points de contours un agent reconnaisseur se déplace aléatoirement, ou dans la direction qui lui a été communiquée par l'agent coordonnateur suite à la prédiction de segments par ce dernier. Pour chaque couple de segments parcourus par un agent reconnaisseur, ce dernier procède à l'appariement de l'objet associé, formé par l'ensemble des segments parcourus par l'agent, aux différents modèles candidats. Pour chaque correspondance, le modèle associé est mémorisé au niveau de l'agent reconnaisseur et l'hypothèse de correspondance est communiquée à l'agent coordonnateur. La mise en correspondance objet-modèle consiste en la comparaison des segments de l'objet, exprimés par rapport au référentiel absolu, aux segments du modèle, exprimés par rapport à la base relative au modèle (*BRM*), cette dernière est calculée de telle sorte que l'expression de deux segments (exp  $S_1$  et  $S_2$  dans la figure IV.9) par rapport à la base absolue ( $o, i, j, k$ ) soit égale à l'expression des deux segments ( $q_1$  et  $q_2$ ) par rapport à cette base ( $p, r, u, v$ ).



**Figure IV.9** – Génération d'une hypothèse de correspondance objet-modèle

Une hypothèse de correspondance objet-modèle est retenue et communiquée au agent coordonnateur si tous les segments parcourus par l'agent reconnaisseur sont superposés, même partiellement, aux segments du modèle ; projeté dans son référentiel relatif.

### 2.3.2 L'agent coordonnateur :

L'agent coordonnateur assure la cohérence des différentes hypothèses de correspondance objet-modèle qui ont été établies par les agents reconnaisseurs. L'évolution des clusters d'agents reconnaisseurs, qui correspondent aux différents objets de la scène, tend à se stabiliser, quand les clusters deviennent constants. A la stabilité des clusters, chacun contient les agents qui parcourent les segments d'un même objet. Le meilleur cluster auquel un agent libre sera affecté est celui qui correspond à la distance minimale entre son centre de gravité, calculé par les positions des différents agents reconnaisseurs formant le cluster, et la position de l'agent libre en question. Une fois le cluster sélectionné, le segment du modèle, projeté dans la *BRM*, le plus proche de l'agent représente le segment vers lequel l'agent libre sera dirigé.

Chaque hypothèse de correspondance objet-modèle consiste en un couple  $(M, Pos)$  du modèle  $M$  correspondant à l'objet, et de la position  $Pos$  de ce dernier par rapport au modèle. Cette dernière est exprimée simplement par la *BRM* qui a permis de générer l'hypothèse de correspondance. A la réception de toute nouvelle hypothèse de correspondance, l'agent coordonnateur procède à l'élimination de certaines hypothèses qui sont incompatible avec l'hypothèse récemment reçue. Une hypothèse stockée  $(M', Pos')$  est éliminée par l'hypothèse reçue  $(M, Pos)$  si les deux positions  $Pos$  et  $Pos'$  désignent la même position spatiale, et le modèle  $M$  contient au moins un segment parcouru, qui ne correspond à aucun segment dans le modèle  $M'$ . L'élimination d'une hypothèse implique l'élimination de l'agent qui l'a créé du cluster correspondant.

Dans la figure [IV.10a](#) les agents  $A_1$  et  $A_2$  établissent les hypothèse  $(M_1, Pos)$  et  $(M_2, Pos)$ . L'objet parcouru peut être l'un des deux Modèles  $M_1$  ou  $M_2$  (Figure [IV.10c](#))

$$A_1 \rightarrow \{(M_1, Pos), (M_2, Pos)\}$$

$$A_2 \rightarrow \{(M_1, Pos), (M_2, Pos)\}$$

A ce stade, aucune des hypothèses précédentes n'est éliminée parce qu'elles sont simultanément compatibles. Après que chaque agent passe par son sommet correspondant (Figure IV.10b), les hypothèses communiquées à l'agent coordonnateur sont les suivantes :

$$A_1 \rightarrow \{(M_1, Pos), (M_2, Pos)\}$$

$$A_2 \rightarrow \{(M_1, Pos)\}$$

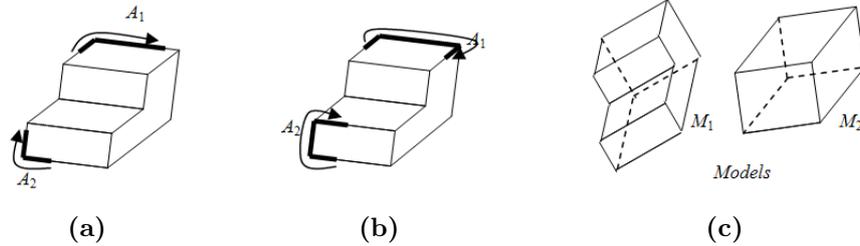


Figure IV.10 – Hypothèses de correspondance objet-modèle (a,b) ; (c) Modèles d'objets

### 3 Conclusion

Dans ce chapitre, nous avons proposé une approche bayésienne à base d'agents pour la reconnaissance d'objets polyédriques à partir d'images de profondeur. En premier temps une population d'agents réactifs évoluent sur l'image pour extraire des segments de droites et les coder d'une structure de données. Un agent bayésien qui est situé au deuxième niveau de notre système, exploite les connaissances que nous avons sur la nature des scènes polyédriques et faire reconstruction de tous les segments qui ont été extraits par les agents détecteurs afin d'obtenir une modélisation complète d'objets. Une fois la reconstruction bayésienne terminée, des agents reconnaisseurs prennent la tâche du suivi des contours et engendrent des hypothèses de mise en correspondance objets-modèles.

---

---

# Chapitre V

---

## Résultats expérimentaux

### 1 Introduction

Dans ce chapitre nous procédons à la validation de notre approche, cette validation est appliquée sur les trois niveaux du système de façon indépendante. Comme nous l'avons vue dans le chapitre IV, dans le premier niveau nous avons lancé un nombre d'agents détecteurs  $N_{agent}$  et en spécifiant leurs paramètres de détection  $(\Theta, R)$ , quand le système terminera l'exécution on obtient des résultats expérimentaux démontrant le processus d'extraction selon la configuration des paramètres de détection. Dans le deuxième niveau, un autre agent dit agent bayésien parcourt le tableau  $T$  des segments détectés par l'agent détecteur, et pour chaque couple de segments  $(S_t, S_{t'})$  il génère des hypothèses concernant les actions à exécuter sur ce couple, puis il détermine l'hypothèse à adopter. Alors cet agent exploite deux réseaux bayésiens (génération et vérification d'hypothèse) qui sont modélisés par l'outil BayesServer 3.0.

### 2 L'environnement de l'expérimentation

#### 2.1 Le langage de programmation

Nous avons implémenté notre système de reconnaissance en langage de programmation Microsoft Visual Basic 2010. Ce dernier a été installé sur une machine Intel Core2Duo équipée par un système d'exploitation Microsoft Windows 7. En effet, notre choix de ce langage a été justifié par les caractéristiques suivantes :

- C'est un langage très adapté à la programmation Orienté objets, ce qui nous a permis de réaliser de façon facile des systèmes multi-agents.
- Présence d'une bibliothèque riche en composants de traitement d'image.

- S’adaptation avec la bibliothèque Bayes Server qui nous a permit la construction et l’inférence dans les réseaux bayésiens.
- Sa capacité de compréhension de la structure de l’application à l’aide d’une fenêtre qui permet de visualiser le diagramme de classes.

### 2.2 Modélisation Bayésienne(BayServer)

Pour bien modéliser nos réseaux bayésiens nous avons choisi l’outil BayesServer 3.0. Ce dernier est une librairie de classes facile à intégrer dans le langage Microsoft visuel basic 2010 et permet de créer et d’inférer simplement dans les reseaux bayésiens et réseaux bayésiens dynamiques. Voici quelques exemples des classes de cette librairie :

- *Network* : représente la classe réseau baysien.
- *Node* : représente la classe nœud avec un ou plusieurs variables.
- *Link* : représente la classe lien directe entre deux nœuds.
- *Inference* : représente une interface algorithmique de l’inférence.

## 3 La description des agents du système

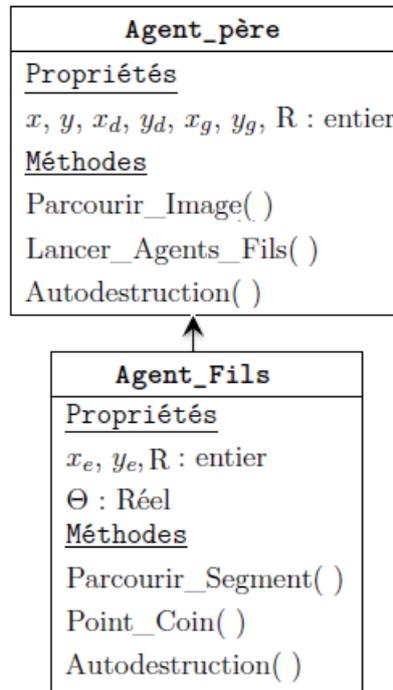
Comme il est indiqué précédemment, notre système est composé de trois niveaux dans lesquels les agents opèrent.

### 3.1 Agent détecteur

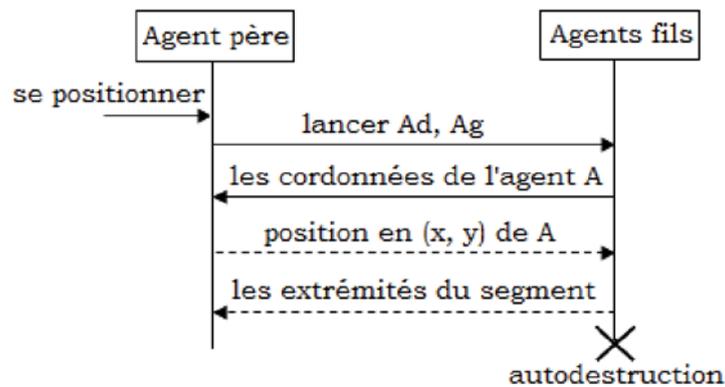
Ce type d’agent fonctionne dans le premier niveau de notre système, et prend la tâche de suivre les points de contour afin de former des segments de droites. En effet, cette population d’agents est composée de deux types d’agents qui sont :

- Agent Père : Il est implémenté par une classe d’objets nommé *Agent\_Père*.
- Agent fils : Il existe en deux instances : agents fils droites et agents fils gauches, qui sont hérités de la classe *Agent\_Père* et nommés *Agents\_Fils*.

La figure V.1 ci-dessous décrire le diagramme de classes et de séquences pour l’agent détecteur :



(a)



(b)

**Figure V.1** – Les diagrammes : (a) Diagramme de classes ; (b) Diagramme de séquences.

Comme validation de l’approche proposée, nous allons exposer et discuter quelques résultats expérimentaux obtenus par le premier niveau du système. Ici le système a été testé sur des images de synthèses de taille 200\*200 pixels, comme il est représenté dans la figure V.2a.



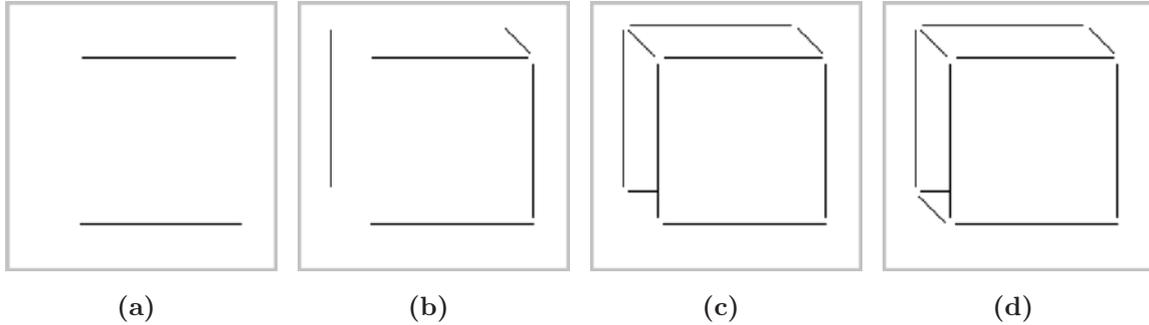
**Figure V.2** – Image de synthèse : (a) Une image représente un cube ; (b) Position d’agents sur le cube .

A chaque fois, nous avons lancé un nombre d’agents donnés  $N_{agent}$  et en spécifiant leurs paramètres de détection  $(\Theta, R)$ . A la fin de d’exécution du système, chaque agent représente un segment de l’objet. Le tableau V.1 ci-dessous montre quelques résultats obtenus en considérant  $R=2$  et  $\theta = 20^\circ$  .

Nombre d’agents lancés	Nombre d’agents détruits	Nombre de segments détectés
1	0	1
7	1	6
10	1	9
25	16	9
26	16	10
>26	>16	10

**Tableau V.1** – Le nombre de segments détectés selon le nombre d’agents lancés (Figure V.1)

Si nous prenons le cas où nous avons lancé 10 agents avec les paramètres de détection  $(\theta = 20^\circ, R = 2)$  sur l’image de synthèse illustré sur la figure V.2a, on obtient les résultats suivants :



**Figure V.3** – Nombre de segments détectés selon le nombre d’agents lancés (NBagent,NBseg) : (a) (NBagent=2 ; NBseg=2) ; (b) (NBagent=5 ; NBseg=5) ; (c) (NBagent=10 ; NBseg=9) ; (d) (NBagent=26 ; NBseg=10).

Le tableau V.2 ci-dessous montre les coordonnées des segments extraits par les agents détecteurs.

N°Segment	$x_1, y_1$	$x_2, y_2$
Segment 1	(149,57)	(65,57)
Segment 2	(152,149)	(64,149)
Segment 3	(153,145)	(153,61)
Segment 4	(151,54)	(138,41)
Segment 5	(42,128)	(42,40)
Segment 6	(57,131)	(45,313)
Segment 7	(61,145)	(61,61)
Segment 8	(134,39)	(46,39)
Segment 9	(59,56)	(44,41)
Segment 10	(58,148)	(44,134)

**Tableau V.2** – Segments trouvés et leurs positions

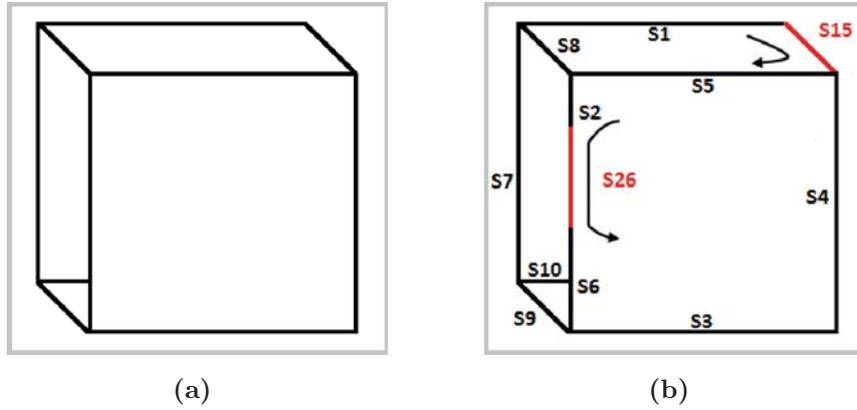
### 3.2 Agent Bayésien :

Cet agent existe en une seule instance et opère de façon autonome dans le deuxième niveau de notre système. En effet, ce dernier a été implémenté avec une classe d’objets nommée *Agent\_bayésien*. Cette classe a comme propriétés (la table des segments  $T$  et les deux réseaux bayésiens : réseau de génération d’hypothèses RGH et réseau de vérification d’hypothèses RVH) et comme méthodes (générer\_hypothèses et vérifier\_hypothèses).

### V.3 La description des agents du système

Nous avons choisi des images de synthèses pour expérimenter le deuxième niveau de notre approche. Nous démontrons ici les résultats issus d'une scène de synthèse comportant un objet cube (Figure V.4a). Cette scène permet d'illustrer l'intérêt de l'agent bayésien.

Dans la figure V.4b nous avons numéroté les segments dans l'ordre de leur détection par les agents détecteurs, et colorer les parties qui présentent les discontinuités sur les segments du cube que nous pouvons traiter par l'agent reconstructeur (agent bayésien).



**Figure V.4** – Image de synthèse : (a) Une image représente un cube ; (b) Un cube présente des discontinuités marquées en rouge .

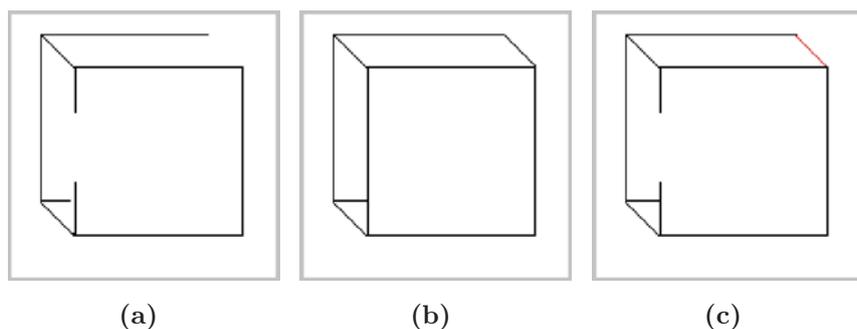
Quand l'agent bayésien parcourt le tableau des segments  $T$  issus de l'image montré dans la figure V.5a, ce dernier observe que les deux segments  $S_1$  et  $S_5$  sont coplanaires et parallèles (ici la coplanarité et le parallélisme est calculés selon un seuil fixé), alors la classe d'hypothèse à laquelle appartiennent ces deux segments est donc la classe  $C3$  ( $C \neg VP$ ), alors l'agent utilise le réseau bayésien génération d'hypothèses pour déterminer l'action à exécuter :

- la probabilité qu'il ya un segment entre  $S_1$  et  $S_5$  est 5%
- la probabilité qu'il ya une jonction entre  $S_1$  et  $S_5$  est 5%
- la probabilité qu'il ya un chainage entre  $S_1$  et  $S_5$  est 85%
- la probabilité qu'il n'ya pas une relation entre  $S_1$  et  $S_5$  est 5%

Conformément à la stratégie codée dans les tables IV.1, IV.2 de probabilité conditionnelle l'hypothèses. Un chainage entre  $S_1$  et  $S_5$  est le plus favorisé.

### V.3 La description des agents du système

Dans la phase de vérification l'agent observe l'image montré dans la figure V.5b qui a été acquise précédemment dans la séquence d'images puis selon son réseau de vérification il valide l'hypothèse principale "Chainage avec un seul segment nommé  $S_{15}$  (Figure V.5c)" avec une probabilité de 85%.



**Figure V.5** – Reconstruction de l'image : (a) Segments détectés ; (b) Image acquise précédemment ; (c) Reconstruction de l'image avec l'ajout d'un segment  $S_{15}$ .

L'objet de la figure V.5c reste toujours incomplet parce qu'il présente une discontinuité entre  $S_2$  et  $S_6$ . Alors l'agent bayésien va examiner les relations entre les deux segments et génère l'hypothèse de l'action à réaliser. L'examen des segments  $S_2$  et  $S_6$  nous a donné la relation suivante :  $S_2$  et  $S_6$  sont voisins coplanaires et non parallèles. Il s'agit donc de la classe  $C2$  ( $CV \neg P$ ).

- la probabilité qu'il y ait un segment entre  $S_2$  et  $S_6$  est 2.5%
- la probabilité qu'il y ait une jonction entre  $S_2$  et  $S_6$  est 60%
- la probabilité qu'il y ait un chainage entre  $S_2$  et  $S_6$  est 30%
- la probabilité qu'il n'y ait pas une relation entre  $S_2$  et  $S_6$  est 7.5%

Conformément à la stratégie codée dans les tables IV.1, IV.2 de probabilité conditionnelle, l'hypothèse d'une jonction entre  $S_2$  et  $S_6$  est la plus favorisée.

Dans la phase de vérification nous observons que la jonction entre  $S_2$  et  $S_6$  n'apparaît dans aucune image acquise précédemment. Afin de vérifier l'hypothèse dans ce cas. Il est nécessaire d'avoir d'autres acquisitions permettant de la vérifier. Dans notre cas, comme il s'agit de la meilleure hypothèse, on l'adopte sans vérification.

## 4 Conclusion

Dans ce chapitre, nous avons exposé quelques résultats expérimentaux de notre approche. En effet, ce chapitre est divisé en deux parties ; dans la première nous avons montré les résultats du premier niveau de notre système sur des images de synthèse. Dans la deuxième partie nous avons discuté les deux cas des résultats issus de l'agent bayésien. Les résultats ainsi présentés montrent le potentiel de notre approche pour l'amélioration à la fois de détection et de reconstruction d'objets polyédriques.

---

## Conclusion générale

Le travail que nous avons présenté le long de ce mémoire apporte sa contribution au problème de la reconnaissance des objets 3D visuels. L'approche proposée ici s'est limitée à des scènes constituées d'objets polyédriques, mais avec ambition future la reconnaissance des scènes de plus en plus complexes.

Le système de reconnaissance qui a été développé, est totalement distribué en utilisant un système multi-agents. Nous nous sommes penchés sur les systèmes multi-agents, dont nous avons proposé trois types d'agents qui opèrent dans une architecture en trois niveaux, respectivement niveau détection de segments de droites, et niveau reconstruction et enfin niveau reconnaissance d'objets.

Au niveau de *détection*, nous avons utilisé une population d'agents évoluant sur des images de profondeur déjà segmentées et suivent les points de contour afin de former des segments de droites, et les enregistrer dans une structure de données.

Nous avons abordé aussi le problème de reconstruction (niveau *reconstruction*) lorsque on a des incertitudes sur les informations visuelles. Pour ce la nous avons proposé un agent dit *agent bayésien* qui exploite les connaissances *a priori* sur les scènes composées d'objets polyédriques (ces connaissances sont codées sur des réseaux bayésiens) afin de former des objets complexes à partir des segments détectés par l'*agent détecteurs*.

L'*agent bayésien* est reposé sur une approche de prédiction/vérification d'hypothèses gérées à l'aide de réseaux bayésiens. Cette approche permet d'obtenir une représentation fiable des objets.

---

Dans le niveau reconnaissance, nous nous sommes basés sur la technique d'appariement Objets-Modèles. Deux types d'agents ont été définis. Le premier représente les *agents reconnaisseurs* qui établissent les hypothèses de correspondance. Le second type d'agents existe en une seule instance et représente l'agent coordonnateur qui examine et valide les différentes hypothèses de mise en correspondance, générées et envoyées par les *agents reconnaisseurs*.

Un certain nombre de questions relatives au travail effectué reste à explorer. Nous en proposons ici quelques unes qui constituent des voies d'investigation intéressantes dans la suite du travail présenté. Il serait notamment souhaitable d'étudier les points suivants :

- La prise en compte d'objets plus complexes que les objets polyédriques.
- Fusionner les deux niveaux détection/reconstruction dans un seul niveau, où les connaissances et les faits sont répartis sur les *agents détecteurs*.

---

## Références bibliographiques

- [AA93] F. ARMAN et J. K. AGGARWAL : Model-based object recognition in dense-range images - a review. *ACM Comput. Surv.*, 25(1):5–43, 1993. 7
- [BB82] D. H. BALLARD et C. M. BROWN : *Computer Vision*. Prentice-Hall, Englewood Cliffs, NJ, 1982. 2
- [BG95] H. BUXTON et S.G. GONG : Visual surveillance in a dynamic and uncertain world. *AI*, 78(1-2):431–459, October 1995. 31
- [CBF03] C. CARABELEA, O. BOISSIER et A. FLOREA : Autonomy in multi-agent systems : A classification attempt. *In Agents and Computational Autonomy*, pages 103–113, 2003. 11
- [CC05] Y. W. CHEN et Y. Q. CHEN : An evolutionary system and its application to automatic image segmentation. *In ICNC (2)*, pages 705–709, 2005. 19
- [CM81] A. CORNUÉJOLS et L. MICLET : Apprentissage artificiel. *In Concepts et algorithmes*, pages 364–365. Eyrolles, 1981. 23, 24
- [Dro94] A. DROGOU : *De la simulation multi-agents à la résolution collective de problèmes*. Thèse de doctorat, Université Paris VI, France, 1994. 12
- [DRP96] D. DJIAN, P. RIVES et P. PROBERT : Reconnaissance de modèles géométriques simples à l'aide de réseaux bayesiens. *In 10ième Congrès AFCET RFIA*, Rennes, France, Janvier 1996. 26
- [Fer95] J. FERBER : *Les Systèmes Multiagents : Vers une Intelligence Collective*. InterEditions, 1995. 11
- [Fer99] J. FERBER : *Multi-Agent Systems : An Introduction to Distributed Artificial Intelligence*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999. 15

- [FG96] S. FRANKLIN et A. C. GRAESSER : Is it an agent, or just a program ? : A taxonomy for autonomous agents. *In ATAL*, pages 21–35, 1996. [10](#)
- [GK94] M. R. GENESERETH et S. P. KETCHPEL : Software agents. *Commun. ACM*, 37(7):48–53, 1994. [10](#)
- [GKG06] S. GHOSH, M. KOTHARI et A. GHOSH : Aggregation pheromone density based image segmentation. *In ICVGIP*, pages 118–127, 2006. [19](#)
- [HM81] R. A. HOWARD et J. E. MATHESON : Influence diagrams. *In R. A. HOWARD et J MATHESON, éditeurs : The Principles and Applications of Decision Analysis*, pages 720–762, CA, 1981. Strategic Decisions Group. [22](#)
- [HM95] R. P. HORAUD et O. MONGA : *Vision par ordinateur : outils fondamentaux*. Editions Hermès, Paris, France., 1995. [7](#), [33](#)
- [HTRC02] F. HARROUET, J. TISSEAU, P. REIGNIER et P. CHEVAILLIER : oris : un environnement de simulation interactive multi-agents. *Technique et Science Informatiques*, 21(4):499–524, 2002. [15](#)
- [LT99] J. LIU et Y. Y. TANG : Adaptive image segmentation with distributed behavior-based agents. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(6):544–551, 1999. [12](#)
- [MC98] É. MARCHAND et F. CHAUMETTE : A bayes nets-based prediction/verification scheme for active visual reconstruction. *In ACCV (1)*, pages 648–655, 1998. [41](#)
- [MGM09] S. MAZOUZI, Z. GUESSOUM et F. MICHEL : Une approche multi-agent pour la segmentation d’images de profondeur à base d’objets polyédriques. une nouvelle approche de segmentation d’images. *Technique et Science Informatiques*, 28(3):365–393, 2009. [33](#)
- [Nat96] E. NATONEK : *3D-MBA : système de reconnaissance d’objets par vision basé modèles 3-D appliqué aux mondes virtuels robotisés*. Thèse de doctorat, Ecole polytechnique fédérale de Lausanne EPFL, Lausanne, 1996. [1](#), [2](#), [3](#), [4](#), [5](#), [6](#)
- [OB02] S. OUADFEL et M. BATOCHE : Ant colony system for image segmentation. *Dans ANTS, LNCS Conferences, Brussels.*, pages 294–295, 2002. [19](#)
- [Pea88] J. PEARL : *Probabilistic reasoning in intelligent systems : networks of plausible inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988. [22](#)
- [RB94] R. D. RIMEY et C. M. BROWN : Control of selective perception using bayes nets and decision theory. *International Journal of Computer Vision*, 12(2-3):173–207, 1994. [30](#)

- [RBG04] V. RODIN, A. BENZINOU et A. GUILLAUD : An immune oriented multi-agent system for biological image processing. *Pattern Recognition*, 37(4):631–645, 2004. [15](#)
- [RTB98] P. REMAGNINO, T. TAN et K. BAKER : Multi-agent visual surveillance of dynamic scenes. *Image Vision Comput.*, 16(8):529–532, 1998. [17](#)
- [Rus97] S. J. RUSSELL : Rationality and intelligence. *Artificial Intelligence*, 94(1-2):57–77, 1997. [10](#)
- [SC94] J. STUTZ et P. CHEESEMAN : A short exposition on bayesian inference and probability. National Aeronautic and Space Administration Ames Research Centre : Computational Sciences Division, Data Learning Group, Juin 1994. [23](#)
- [Sho93] Y. SHOHAM : Agent-oriented programming. *Artificial Intelligence*, 60(1):51–92, 1993. [10](#)
- [Ste03] T. A. STEPHENSON : An introduction to bayesian network theory and usage, Avril 2003. [25](#)
- [Tso88] J. K. TSOTSOS : A 'complexity level' analysis of immediate vision. *International Journal of Computer Vision*, 1(4):303–320, janvier 1988. [2](#)
- [WJ95] M. WOOLDRIDGE et N. R. JENNINGS : Intelligent agents : Theory and practice. *The Knowledge Engineering Review*, 10(2):115–152, 1995. [9](#)
- [Woo02] M. WOOLDRIDGE : *Introduction to multiagent systems*. John Wiley and Sons, 2002. [9](#)
- [ZJW03] F. ZAMBONELLI, N. R. JENNINGS et M. WOOLDRIDGE : Developing multiagent systems : The gaia methodology. *ACM Trans. Softw. Eng. Methodol.*, 12(3):317–370, 2003. [10](#)