HAL
open science

Une architecture multi-agent pour la détection, la reconnaissance et l'identification de cibles

Fabienne Ealet

To cite this version:

Fabienne Ealet. Une architecture multi-agent pour la détection, la reconnaissance et l'identification de cibles. Mathématiques [math]. Université Joseph-Fourier - Grenoble I, 2001. Français. NNT : . tel-00004678

HAL Id: tel-00004678

https://theses.hal.science/tel-00004678v1

Submitted on 16 Feb 2004

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

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

UNIVERSITE JOSEPH FOURIER - Grenoble 1
SCIENCES & GEOGRAPHIE

THESE

présentée pour obtenir

Le GRADE de DOCTEUR EN SCIENCES

discipline : informatique

par

Fabienne EALET

UNE ARCHITECTURE MULTI-AGENT POUR LA DETECTION, LA RECONNAISSANCE ET L'IDENTIFICATION DE CIBLES

soutenue le 25/06/01 devant la Commission d'examen

Alain	Chehikian	Président du Jury
Monique	Thonnat	Rapporteur
Michèle	Rombaut	Rapporteur
Nicole	Cambou	Examinateur
Bertrand	Collin	Examinateur
Georges	Stamon	Examinateur
Catherine	Garbay	Directeur de thèse

Remerciements

Cette thèse CIFRE s'est déroulée au sein du service "géographie imagerie perception"de la DGA (CTA/GIP) en collaboration avec le service vision de EADS, et l´équipe SIC de TIMC-IMAG.

Je tiens en premier lieu à adresser mes sincères remerciements à Madame Catherine Garbay, directeur de recherche au CNRS, qui a accepté de diriger cette thèse. Ses compétences, son exigence, sa rigueur et ses conseils m'ont été d'une grande aide et m'ont permis de progresser au cours de ces trois années. Très sincèrement je souhaite à tous les futurs thésards de pouvoir bénéficier d'un tel encadrement et d'un tel soutient, qu'ils soient scientifique ou humain.

Mes remerciements s'adressent également à Monsieur Bertrand collin, professeur aggrégé de l´école normale supérieure de Cachan à qui, au delà du soutien et de la confiance qu'il m'a témoigné dans le cadre de ce travail, je dois ma formation en traitement d'images et en informatique.

Je remercie Nicole Cambou, chef du service vision à EADS, pour sa confiance et les libertés qu'elle m'a accordées.

J'adresse toute ma gratitude à Monsieur Chehikian, professeur à l'université Joseph Fourier de Grenoble pour m'avoir fait l'honneur de présider ce jury ainsi que Madame Michèle Rombaut et Madame Monique Thonnat pour avoir accepté la lourde de tâche de juger mon travail en qualité de rapporteur.

Enfin je remercie Monsieur Georges Stamon, d'avoir participé à mon jury. En tant qu'examinateur j'ai particulièrement apprécié ses remarques toujours judicieuses, ses encouragements constants et sa gentillesse.

Je ne saurai oublier dans mes remerciements Monsieur Dominique Luzeaux qui m'a accueilli dans son service après mon stage de DEA, Madame Véronique Serfaty-Dutronc pour ses conseils précieux et son aide permanente. Bien évidemment j'ai une pensée pour tous les thésards Fabrice, Lui, Ronan, Alexis, André, Jérôme, Thomas, Karine, Philippe avec qui j'ai pu faire un bout de route toujours dans la bonne humeur. Merci pour tous ces moments de bonheur et de convivialité. Un grand merci à toute ma famille Loic, mes parents et mes soeurs pour leur patience et leur compréhension.

Table des matières

1 Introduction **13**
1.1 Définition de la fonction DRI 14
1.2 Le problème 16
1.2.1 Notre cadre d'analyse 16
1.2.2 Les missions 17
1.2.3 Les cibles 17
1.3 Les autres données du problème 19
1.4 Un système de vision pour la DRI 19
1.5 Plan de lecture 20

2 État de l'art de la fonction DRI **23**
2.1 Introduction 25
2.2 Les modèles 25
2.2.1 Les modèles de scène 27
2.2.2 Les modèles d'observations 31
2.2.3 Conclusion 36
2.3 Les techniques de mise en correspondance 38
2.3.1 La théorie des graphes 38
2.3.2 Les techniques de seuillage 39
2.3.3 Les calculs de distance 39
2.3.4 Les classifieurs 40
2.3.5 Conclusion 42
2.4 Processus de reconnaissance 42
2.4.1 La relaxation 43
2.4.2 la recherche arborescente 43
2.4.3 La prédiction et vérification d'hypothèses 43
2.4.4 La fonction coût/bénéfice 46
2.4.5 Conclusion 47
2.5 Conclusion 48

3 L'approche proposée **51**
3.1 Introduction 53
3.2 Vers un système de vision 53
3.2.1 Les contraintes 54
3.2.2 L'approche proposée 54

3.3 Focalisation et adaptation 57
3.4 Distribution et coopération 59
3.4.1 Coopération confrontative 60
3.4.2 Coopération augmentative 61
3.4.3 Coopération intégrative 61
3.4.4 Conclusion 62
3.5 L'incertitude 62
3.5.1 Les réseaux bayésiens 63
3.5.2 Les chaines de Markov cachées 66
3.5.3 En conclusion 68
3.6 Incrémentalité et apprentissage 68
3.7 Conclusion 69

4 Description du système 71
4.1 Introduction 73
4.2 Principe de fonctionnement 73
4.3 Plate forme multi-agent 75
4.3.1 Présentation de la plate-forme 75
4.3.2 Le modèle d'agent 76
4.3.3 L'administrateur 77
4.4 Description fonctionnelle du système 77
4.4.1 Les notations adoptées 77
4.4.2 Le cycle d'analyse intra-image 78
4.4.3 Le cycle d'analyse inter-image. 81
4.5 Représentation des informations 83
4.5.1 La base de connaissances 83
4.5.2 Le modèle du monde 89
4.6 Comportements des agents 92
4.6.1 L'agent concept 92
4.6.2 L'agent localisation 94
4.6.3 L'agent détection 99
4.6.4 L'agent focalisation 102
4.6.5 L'agent mouvement 103
4.6.6 L'agent Décision 104
4.7 Conclusion 106

5 Résultats expérimentaux 109
5.1 Introduction 111
5.2 Quelques caractéristiques techniques 111
5.3 Présentation du cadre d'application 111
5.4 Exploration spatiale de l'image 114
5.4.1 Analyse directe 115
5.4.2 Apport du contexte 119
5.5 Exploitation d'une séquence d'images 127
5.5.1 Les capacités d'adaptation du système 128

5.5.2 L'adaptation de la population d'agents 130
5.5.3 Adaptation de la base de connaissance et du modèle du monde 133
5.5.4 Influence d'une perturbation 134
5.5.5 Influence du seuil de validation 138
5.6 Conclusion 139

6 Vers la reconnaissance des cibles 143
6.1 Introduction 145
6.2 Présentation du système de reconnaissance 145
6.3 Description des agents 147
6.3.1 L'agent Position 147
6.3.2 L'agent Identité 148
6.4 Les attributs pour la reconnaissance 148
6.4.1 Mesures géométriques 149
6.4.2 Mesures physiques cinématiques 151
6.4.3 Mesures statistiques 151
6.5 Apprentissage 152
6.5.1 Les images 152
6.5.2 Structuration de la base de données 153
6.5.3 L'analyse en composante principale 154
6.5.4 L'Analyse factorielle discriminante (AFD) 157
6.6 La fonction utilité 158
6.6.1 Diagramme d'influence 158
6.6.2 Un exemple de diagramme 159
6.7 Conclusion 160

7 Conclusion 163

A Les réseaux bayésiens : "The bayes Net Toolbox (BNT)" 167
A.1 Exemple de réseau 167
A.2 Construction d'un graphe 168
A.3 L'inférence 169
A.4 Vérification par le calcul 170

B Le modèle d'agent, l'administrateur et les fiches d'informations 171
B.1 Le modèle d'agent 171
B.2 L'administrateur 172
B.3 Les fiches d'informations 173
B.4 Communication entre agents 174

C Algorithmes de traitement d'images 175
C.1 Les méthodes et leurs paramètres 175
C.1.1 Les méthodes de l'agent Localisation 175
C.1.2 Les méthodes des autres agents 176
C.2 La transformée de Hough : 177
C.2.1 Le principe 177

C.2.2 Les méthodes déterministes 178
C.2.3 Les approches probabilistes 179
C.3 Enveloppe convexe : algorithme de Graham 181
C.4 Les mesures d'orientation 181
C.5 Détecteur de Harris 182
C.6 Les ondelettes de Gabor 182

D Exemple : modèle du monde 183

Table des figures

1.1 Exemples de véhicules en phase de détection. 14
1.2 Exemple d'une cible en phase d'identification. 15
1.3 Exemples de cibles militaires 18
1.4 Illustration de la variabilité inter-classe. 18
1.5 Illustration de la variabilité intra-classe. 18

2.1 Schéma de principe de la fonction de détection, de reconnaissance et d'identification de cibles 26
2.2 Description hiérarchique de véhicules militaires. 29
2.3 Graphe de composition et de spécialisation 30
2.4 Récapitulatifs des modélisations géométriques proposées pour réaliser le processus DRI. 35
2.5 Reconnaissance par partie de la cible. 41
2.6 La boucle de rétroaction au sein de la fonction DRI inclut la phase de détection 44
2.7 La boucle de rétroaction au sein de la fonction DRI n'inclut pas la phase de détection. 44
2.8 Schéma de principe du système multi-capteur de reconnaissance automatique de cibles développé par Beveridge *et al.* 45
2.9 Le système de reconnaissance automatique de cibles : ASAAT 46
2.10 Diagramme représentant le système DRI proposé par Roberts et Brown. 47

3.1 Vue générale du système multi-agent. 55
3.2 Approche proposée pour l'analyse de scènes. 57
3.3 Exemple de réseau bayésien 64
3.4 Exemples de chaîne de Markov 66
3.5 Exemples d'un modèle de Markov caché 67
3.6 Représentation d'un modèle de Markov caché sous forme d'un réseau bayésien 68

4.1 Les concepts de la scène 74
4.2 L'architecture du système multi-agent. 75
4.3 Description de la plate-forme multi-agent 76
4.4 Les structures issues de l'image 78
4.5 Exploration spatiale de l'image 80
4.6 Le schéma de contrôle intra-image 82

4.7 La base de connaissances 85
4.8 Base de connaissances : description des liens explicites entre les espaces de représentation 86
4.9 Premier module de l'interface graphique. 87
4.10 Le second module de l'interface graphique 88
4.11 Troisième module de l'interface graphique permettant de générer la base de connaissances. 88
4.12 Structure du modèle du monde 90
4.13 Description du réseau bayésien permettant le calcul de la confiance d'un support 92
4.14 Schéma de principe de l'agent Concept 94
4.15 Principe de localisation des échantillons dans une zone d'intérêt . . . 95
4.16 Distributions de probabilités associées au modèle de la base de connaissances 96
4.17 Distributions de probabilités associées aux informations 97
4.18 Calcul de la confiance d'un échantillon 97
4.19 Le calcul de la confiance dans un échantillon 98
4.20 Schéma de principe de l'agent Localisation 99
4.21 Réseau bayésien pour calculer la confiance dans un support 100
4.22 Schéma de principe de l'agent Détection. 101
4.23 Principe de focalisation intra-image 102
4.24 Exemple de stratégie de focalisation. 102
4.25 Schéma de principe de l'agent Focalisation 103
4.26 Schéma de principe de l'agent Mouvement 105
4.27 Comportement de l'agent Décision 106
4.28 Schéma de principe de l'agent Décision 106

5.1 Enchainement des étapes pour l'analyse directe. 117
5.2 Segmentation des supports véhicule 117
5.3 Graphe d'agents pour détecter les véhicules 118
5.4 Répartition des temps de calculs entre les différents agents 119
5.5 Autre segmentation de supports véhicule 120
5.6 Image représentant les zones d'intérêt associées à l'objectif véhicule . 122
5.7 Réseau bayésien utilisé pour calculer la confiance dans un support. . . 123
5.8 Segmentation des véhicules avec une règle d'exploration spatiale. . . . 124
5.9 Graphe d'agents réalisant la détection de véhicules suivant la stratégie d'exploration "véhicules sur routes". 124
5.10 Exemple de fichier de configuration 126
5.11 Ensemble des zones d'intérêt définies par les agents Focalisation . . . 127
5.12 Segmentation des véhicules avec deux stratégies de focalisation 127
5.13 Graphe d'agents relatif à la détection de véhicules lorsque deux règles de focalisation sont spécifiées. 128
5.14 Séquence de référence 129
5.15 Fichier de configuration associé à la séquence de référence 131
5.16 Segmentation des zones en mouvement 131

5.17 Graphe d'agents en présence de mouvement 132
5.18 Répartition des temps de calculs entre agents 133
5.19 Fichier de configuration relatif à la première image de la séquence. . . 134
5.20 Résultat de segmentation 135
5.21 Numérotation des supports d'objets 136
5.22 Description des différents supports présents dans la séquence d'images.137
5.23 Segmentation de la première image avec différents seuils de validation 139

6.1 Différentes vues d'un mirage 2000. 146
6.2 Description schématique du système de reconnaissance. 147
6.3 Image de synthèse générée par le logiciel SPECRAY. 153
6.4 Le référentiel choisi. 153
6.5 L'espace d'observation est composé de groupes de positions. 154
6.6 Interface pour réaliser l'analyse en composantes principales 155
6.7 Analyse en composantes principales obtenue pour les chars vus de face(gris), de profil (en rose) et de face (rouge) 155
6.8 Analyse en composantes principales avec huit attributs pour trois types de cibles vues de profil. 156
6.9 Analyse en composantes principales avec huit attributs pour trois types de cibles vues de face. 157
6.10 Exemple d'utilisation des diagrammes d'influence dans le domaine de la poursuite. 159
6.11 Exemple de diagramme d'influence 160

A.1 Un exemple de réseau bayésien composé de 4 variables aléatoires. . . 168

C.1 La transformée de Hough d'un point 177
C.2 La transformée de Hough d'une droite 178

Chapitre 1

Introduction

Sommaire

1.1 Définition de la fonction DRI **14**
1.2 Le problème **16**
1.3 Les autres données du problème **19**
1.4 Un système de vision pour la DRI **19**
1.5 Plan de lecture **20**

Pour l'homme, l'acquisition des images et l'identification de cibles s'effectuent par l'intermédiaire du système visuel. L'œil peut être défini comme un capteur qui va transmettre ses données au cerveau de façon à traiter les informations et prendre une décision. Le système de perception visuel humain "œil-cerveau" est régi par des mécanismes très complexes, qui ont toutefois des limitations. Par exemple, l'œil n'est sensible qu'à certaines longueurs d'onde, sa sensibilité diminue avec l'obscurité, et le champ visuel est limité. Pour pallier ces problèmes, des systèmes complexes d'aide à la décision ont été développés capable de percevoir l'environnement au-delà du système sensoriel et de réaliser des tâches de plus en plus complexes.

Pour répondre à ce besoin, dans le domaine militaire, différentes technologies ont été mises au point par l'intermédiaire de capteurs spécifiques tels que l'imagerie infrarouge et thermique, le radar, la télémétrie laser ou encore les intensificateurs de lumières. En contrepartie les quantités d'informations à gérer sont devenues gigantesques et délicates voire impossibles à traiter rapidement pour prendre une décision.

Les besoins en matière d'aide à la décision requièrent dans le domaine militaire, l'élaboration d'informations toujours plus précises, synthétiques et fiables et ceci dans le cadre de conditions opérationnelles souvent difficiles, avec des délais de réaction de plus en plus courts. Le but n'est plus seulement d'accumuler des informations, il faut impérativement les analyser et les interpréter de façon à prendre

une décision rapide qui soit adaptée à la situation.

C'est dans ce cadre que vient s'inscrire ce travail. L'idée directrice est de proposer une architecture pour l'analyse d'images s'articulant, d'une part autour de la modélisation de la cible, et d'autre part autour de la mise en œuvre d'une chaîne d'opérateurs réalisant la fonction Détection - Reconnaissance - Identification (DRI) de cibles.

Pour suivre l'ordre chronologique de la démarche de recherche que nous avons adoptée, nous poserons dans un premier temps le problème de la DRI, puis nous en préciserons le domaine d'application ainsi que les données et les contraintes opérationnelles.

1.1 Définition de la fonction DRI

Le processus de DRI doit être perçu comme un enrichissement temporel des informations délivrées par le système, et relatives à la cible notamment.

▽ La phase de détection :

Classiquement, la phase de **détection** s'établit à très grande distance. Elle ne dispose que de peu d'information : la cible est représentée par une dizaine de pixels environ. Un exemple est présenté sur la figure 1.1. Les flèches pointent les véhicules présents dans l'image.

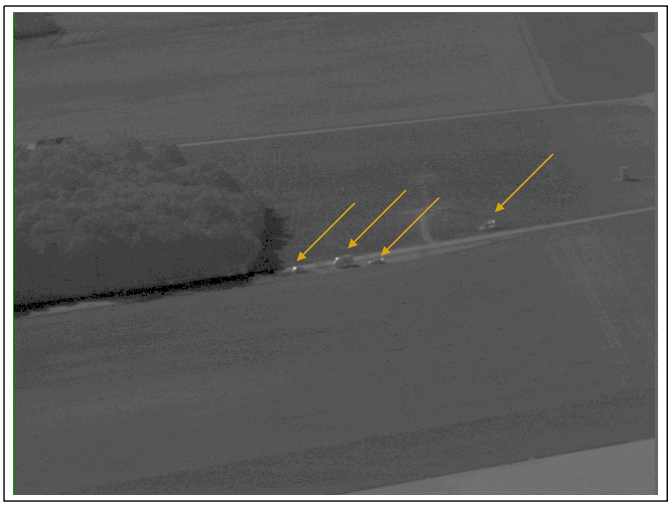


FIG. 1.1 – Exemples de véhicules en phase de détection. Chaque cible est pointée par une flèche.

L'objectif est de trouver l'information cible dans un environnement image caractérisé par une complexité plus ou moins importante. Cette étape détecte des

cibles dans l'espace considéré sans pour autant en préciser ni la classe, ni le nom. Sur la figure 1.1, nous apercevons des véhicules, une route, un bosquet et enfin des champs. Une analyse plus attentive nous permet de remarquer un bosquet dans le coin supérieur gauche de l'image, et une seconde route dans l'image. À ce niveau, on peut déjà donner une interprétation et une description symbolique de la scène.

Dans l'image il y a deux routes. Sur la route au centre de l'image se situent des véhicules. À gauche de l'image deux bosquets se distinguent.

▽ La phase de reconnaissance :

L'objectif de la **reconnaissance** est de fournir une description symbolique de la cible et de son environnement. L'observation est affinée. Des paramètres descriptifs permettent de qualifier la cible. Les données image relatives à la cible sont représentées par quelques dizaines de pixels (entre 50 et 150 pixels). Sur la base de cette information, il est possible d'envisager la spécification de la classe de l'objet. Par exemple dans le cas des cibles militaires, on distinguera les avions, les bateaux et les véhicules au sol. Ainsi, sur la figure 1.2, il est possible de définir la classe des cibles. C'est un véhicule au sol, et non un navire ou un avion.

▽ La phase d'identification :

L'**identification** correspond à la dernière phase du processus d'analyse. Elle nécessite une meilleure résolution de l'image et en particulier de la cible. L'information cible est représentée par un minimum de 150 pixels environ. Un exemple est présenté sur la figure 1.2. La cible est encadrée et représente un véhicule blindé.

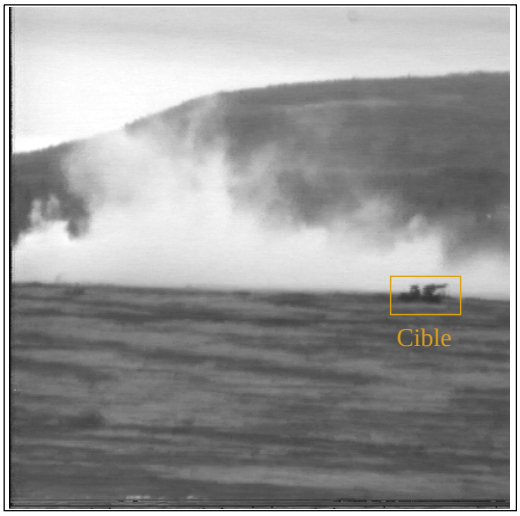


Fig. 1.2 – Exemple d'une cible en phase d'identification. La cible est encadrée.

L'information significative doit permettre de nommer l'objet en question. Elle distingue par conséquent les éléments d'une même classe. Notamment, pour la classe

véhicule, il faut être en mesure de répondre à la question suivante : est-ce un char de type AMX10, AMX30, ou Marder ?

La fonction DRI cherchera en permanence un compromis entre une distance à la cible toujours plus grande (pour augmenter le temps de réaction vis-à-vis d'une menace éventuelle) et une information suffisamment riche pour discriminer tout en minimisant le nombre de fausses alarmes. L'objectif est bien évidemment d'identifier le plus tôt possible d'éventuels dangers.

Le résultat de la fonction DRI est un modèle symbolique composé des différentes cibles présentes, et de leurs positions exactes ou relatives. En phase de détection, cette fonction se décompose en deux étapes. La première étape consiste à trouver les grandes structures de la scène (routes, végétation, fumée...). Les traitements mis en œuvre sont des outils de segmentation. La seconde étape est de contraindre l'analyse, selon divers critères tels que la position, la taille, le contexte etc. Le but recherché est de compenser le manque d'information sur la cible par des connaissances environnementales. Plus la cible se rapproche du capteur, plus les données la caractérisant s'enrichissent et deviennent suffisantes pour reconnaître et puis identifier l'objet. L'importance des connaissances liées au contexte est alors moindre.

1.2 Le problème

Par l'intermédiaire de la DRI, l'objectif actuel est de fournir une aide à la décision pour l'opérationnel. Dans le cadre du domaine militaire, il s'agit donc davantage de tirer profit du système semi-automatique plutôt que de proposer un système automatique.

Il faut être en mesure de prendre le plus rapidement possible une décision, avec le minimum d'imprécision. Un système DRI doit en permanence rechercher l'information pertinente dans une multitude d'informations sans intérêt. Le problème fondamental est la détection, la reconnaissance et l'identification d'objets d'intérêt (les cibles) dans un environnement où d'autres objets (sans intérêt) peuvent être présents.

1.2.1 Notre cadre d'analyse

Le contexte d'analyse de la fonction DRI peut être très varié. Nous avons choisi de restreindre le cadre de cette étude à la recherche de cibles de type militaires dans un environnement extérieur non structuré (non urbain). Dans ce contexte d'étude, il apparaît :

- une grande diversité des scenarii possibles, sachant qu'un scénario correspond à une phase d'observation d'une ou plusieurs cibles dans l'environnement considéré.

- une variabilité au sein d'un même scénario. Il existe une évolution constante en terme de luminosité, d'occultations éventuelles, d'ombres...
- une incomplétude et une imprécision dans les informations fournies par le capteur.

Cette information imparfaite et incomplète donne une vue partielle de la scène. Mais, elle ne fournit pas suffisamment d'éléments pour décrire complètement cette scène. Il manque, par exemple, des attributs de profondeur. D'autre part, il est quasiment impossible de quantifier l'influence d'un élément extérieur en chaque pixel de l'image. L'apparence d'un objet dépend du matériau, des conditions atmosphériques, de l'angle d'observation, de la position des sources de lumière, etc. Toutes ces informations sont confondues et se traduisent par une intensité en chaque pixel de l'image. Ainsi, la compréhension d'une image nécessite des connaissances *a priori* sur le domaine d'application, qui s'effectue dans l'étape de la préparation de mission spécifiant le domaine d'application, le contexte, le type de cibles ainsi que le type et la position des capteurs.

1.2.2 Les missions

Toujours dans le cadre de notre étude, nous distinguons trois types de missions pour effectuer la fonction DRI. Elles sont fonction de la position de la caméra :

- les missions "sol-sol" : le porteur de la caméra est placé au sol et observe une scène au sol. La cible éventuelle évolue dans l'espace observé. Le capteur peut-être fixe ou mobile. Les phénomènes d'occultation sont, dans cette configuration, très fréquents. Ils sont dus notamment à la présence de végétation ;
- les missions "air-sol" : le capteur est embarqué sur un porteur "air" (avion, hélicoptère). Il est mobile et il observe une scène au sol. Les occultations sont dues principalement à des structures masquantes de type végétation.
- les missions "air-air" : le capteur est embarqué sur un porteur "air" (avion, hélicoptère). Il est mobile et il observe un fond de ciel. Les occultations des cibles sont dues principalement à la présence de nuages, ou éventuellement d'autres cibles.

Suivant le type de mission, les objectifs, les classes de cibles et leurs caractéristiques d'évolution varient. Dans une mission sol-air les cibles recherchées sont les avions. Dans une mission air-sol ou sol-sol, l'intérêt est porté davantage sur les bateaux ou les véhicules au sol. Cette spécification donne une information sur la classe éventuelle de la cible.

1.2.3 Les cibles

Nous nous plaçons dans le cadre de la recherche de cibles dont la classe est connue *a priori*. Nous considérerons exclusivement des cibles de type militaire. Nous distinguons trois classes de cibles : les véhicules au sol, les avions et les navires.

Nous donnons ci-dessous (Tab 1.3) un bref aperçu des cibles potentiellement présentes dans les images.

Véhicules	Avions	Navires
AMX 10	Jaguar	Frégate
AMX 30	Mirage	Foch
Marder	Rafale	Croiseur
Roland	Transall	
Shelter	Sukoi	...
...	...	

FIG. 1.3 – Exemples de cibles militaires. Il existe trois classes de cibles : véhicules, avions et navires. Chaque classe est constituée d'un ensemble de cibles

Ce classement traduit une très grande variabilité intra-classe et inter-classe. Comme on peut le constater sur la figure 1.4, il est évident que la signature d'un avion et d'un char sera toujours très différente.

FIG. 1.4 – Illustration de la variabilité inter-classe entre trois cibles : un AMX30 à gauche, le porte-avion Foch au centre et un transall à droite.

Mais il existe également une très grande variabilité d'ordre géométrique et radiométrique entre un Jaguar et un Transall, entre un porte-avion et une frégate. Cette variabilité intra-classe est illustrée pour les véhicules au sol sur la figure 1.5 :

FIG. 1.5 – Illustration de la variabilité intra-classe entre trois véhicules : un AMX10 à gauche, un AMX30 au centre et un char Roland à droite.

Cette variabilité impose un certain nombre de contraintes et en premier lieu la gestion des incertitudes dans le système.

1.3 Les autres données du problème

Un système DRI a pour vocation d'exploiter les "données image" pour permettre dans un mode automatique ou semi-automatique la détection, la reconnaissance et l'identification des cibles. Dans le cadre de notre étude, les données image correspondent à une séquence temporelle d'images d'une scène réelle acquise(s) par un capteur de type infrarouge ou visible.

Pour réaliser ce changement de représentation entre l'espace image et l'interprétation symbolique, il est nécessaire de faire appel à un ensemble très vaste de techniques qui regroupe notamment l'ensemble des techniques relatives aux domaines de la reconnaissance des formes, de la classification, de l'intelligence artificielle et bien entendu les techniques de mises en correspondance [Dudgeon et Lacoss, 1993]. La phase de mise en correspondance s'appuie en général sur des connaissances liées au domaine, et à la caractérisation des objets. Cette connaissance est spécifiée *a priori* dans une ou plusieurs bases de données du système.

Dans cette étude, nous n'avons pas considéré le problème de l'acquisition des images numériques qui contient sa propre panoplie de transformations, de corrections, d'améliorations d'images, et de difficultés... L'ensemble des données capteur est spécifié *a priori* dans la phase de préparation de mission qui constitue la dernière source d'informations.

1.4 Un système de vision pour la DRI

Au cours des deux dernières décennies, les avancées sur le thème de la DRI ont été considérables d'un point de vue scientifique et technique. Bhanu et Ratches en dressent une revue particulièrement riche [Bhanu, 1986; Ratches *et al.*, 1997].

Les systèmes proposés au début des années 1980 étaient pour la plupart heuristiques. Typiquement, ils suivaient pour la plupart le schéma séquentiel suivant : segmentation, détection, extraction de paramètres, classification et suivi de cibles.

À la fin des années 1980, une nouvelle génération de systèmes est apparue, cherchant à briser cette séquentialité. Ces approches intégraient explicitement la connaissance et les techniques de reconnaissance de formes (*knowledge-based vision et model-based vision*). Elles enchaînaient généralement les étapes suivantes : la recherche de zones d'intérêt puis la reconnaissance et l'identification des cibles sur un domaine plus restreint.

Aujourd'hui l'amélioration des performances nécessite l'association d'informations complémentaires afin de donner une réponse adaptée aux besoins opérationnels pour l'analyse de situations. Elle permet notamment d'améliorer la robustesse de l'analyse vis-à-vis de son environnement, l'acuité des informations transmises et la réactivité temporelle. Toutefois l'exploitation pertinente des synergies entre les différentes sources d'informations et le volume des données disponibles requiert la mise en œuvre de méthodes de fusion de données. Différentes approches théoriques sont *a priori* susceptibles de satisfaire cette contrainte puisqu'elles permettent de gérer par exemple l'incertitude, l'imprécision, l'incomplétude, la fiabilité, la dépendance ou la pertinence. Parmi celles-ci il convient de citer principalement les théories probabilistes, possibilistes, des ensembles flous, de l'évidence, les approches logiques ou encore les méthodes connexionnistes.

L'intégration de données multi-capteur, de connaissances spatiales et temporelles ont permis d'ouvrir de nouveaux horizons. Le but n'est plus de développer de nouveaux algorithmes mais davantage de tirer profit de l'information partielle et imprécise fournie par chacun d'eux. Les contraintes liées à l'hétérogénéité et à la quantité des connaissances conduisent à des architectures distribuant la connaissance. L'objectif désormais est de doter le système DRI d'une capacité à gérer, sélectionner et activer ses propres ressources de façon à construire des stratégies de contrôle efficaces pour mener à bien les différentes phases d'analyse. Un système DRI se situe aujourd'hui à la croisée des chemins entre l'analyse d'image et les fonctions cognitives de raisonnement.

Dans le cadre de ce travail, nous proposons une approche s'appuyant sur des concepts multi-agents pour résoudre le problème de la DRI. À chaque instant, une population d'agents travaille en parallèle dans l'environnement. Chaque agent dispose de connaissances descriptives et opératoires lui permettant d'élaborer sa propre stratégie. Chaque agent effectue des traitements et stocke ses résultats dans un modèle du monde commun à tous les agents.

1.5 Plan de lecture

Ce manuscrit est organisé en quatre chapitres :

- le chapitre 2 dresse un rappel concernant les différentes approches mises en œuvre pour la DRI. Cet état de l'art est décliné suivant les modèles, les techniques de mise en correspondance et les processus de reconnaissance utilisés.

- le chapitre 3 élargit le cadre de notre étude à la notion de systèmes de vision et d'interprétation de scène. Le but est de tirer profit de ces architectures, de leur approche de modélisation des connaissances et de gestion de l'incertain, ainsi que des stratégies mises en œuvre.

- le chapitre 4 donne une description complète de notre système en détaillant son architecture, ses agents, les types de connaissances manipulées, les comportements des agents et enfin la méthode de gestion de l'incertitude.

- le chapitre 5 présente les résultats obtenus dans le cadre de notre étude.

- le chapitre 6 propose une extension au processus de reconnaissance de cibles.

Enfin nous terminerons ce manuscrit par une conclusion et une présentation des perspectives.

Chapitre 2

État de l'art de la fonction DRI

Sommaire

2.1 Introduction **25**
2.2 Les modèles **25**
2.3 Les techniques de mise en correspondance **38**
2.4 Processus de reconnaissance **42**
2.5 Conclusion **48**

Résumé

Par définition un modèle est une vue simplifiée d'un monde souvent très complexe et variable dans le temps. Pour réaliser le processus DRI, il convient de choisir une modélisation qui soit représentative de la réalité même s'il existe toujours un décalage entre cette réalité et l'observable, et suffisamment robuste dans le contexte d'application. Nous distinguerons deux classes de modèles : les modèles de scènes qui sont connus a priori et les modèles d'observation qui sont issus des mesures. Nous dresserons dans ce chapitre un panorama des différentes modélisations utilisées.

Pour formuler une hypothèse concernant la présence d'une cible, sa classe ou son identité, il faut être en mesure de mettre en correspondance les modèles de scènes et d'observation. Le but est de calculer une distance entre les observations obtenues et le modèle de l'objet connu a priori. Nous détaillerons les différentes approches mises en œuvre.

Nous terminerons ce chapitre en présentant les processus de reconnaissance au sein des systèmes DRI. La difficulté du problème liée à la gestion des hypothèses incite à développer des mécanismes de remise en cause basés notamment sur des boucles de rétroaction.

2.1 Introduction

Par définition un modèle est une vue simplifiée d'un monde souvent très complexe et variable dans le temps. Aussi, il existe toujours un décalage entre la réalité et le modèle qui donne une représentation de ce monde réel sous forme d'un modèle de scène.

Dans un système de vision et particulièrement dans un système DRI, on ne dispose que d'une information incomplète et imparfaite. Ces imperfections sont dues en partie à l'acquisition des images qui introduit à tous les niveaux un peu plus de déformations. La non complétude introduit également un décalage entre l'observation et le monde physique réel.

La reconnaissance automatique de cibles consiste à mettre en correspondance ces deux modèles imparfaits que sont les modèles d'observations et de scène de manière à caractériser la cible par rapport à son milieu environnant [Djian, 1997]. Une représentation très schématique présentée, sur la figure 2.1, met en évidence cet aspect. La fonction DRI utilise successivement :

- des modèles : les modèles de scène donnent une représentation interne au système de la connaissance sur le monde physique ; les modèles d'observations donnent une représentation des données délivrées par le capteur. Chaque modèle peut être traduit sous des formes diverses basées sur le calcul de primitives ou d'attributs. Les primitives sont extraites de l'image et constituées par un sous-ensemble de points représentant des parties significatives de l'objet ou de la forme. Les primitives les plus couramment utilisées sont les régions et les contours. Les attributs sont des mesures qui sont effectuées soit sur ces primitives précédemment extraites soit sur les données images brutes ;
- des techniques de mise en correspondance. On s'intéressera au calcul de la distance entre ces deux catégories de modèles (modèles de scène et modèle d'observation) ;
- un processus de reconnaissance qui spécifie l'articulation entre les trois étapes que sont la détection, la reconnaissance et l'identification. Il décrit l'enchaînement des étapes et les boucles de rétroaction éventuelles.

Dans les sections suivantes, nous détaillerons chacune de ces trois composantes que sont les modèles, les techniques de mise en correspondance et les processus de reconnaissance.

2.2 Les modèles

Dans le processus DRI, principalement basé sur des techniques de mise en correspondance, il faut comparer deux modèles. Le premier va se focaliser sur ce que l'on connaît *a priori*, alors que le second sera issu d'une mesure. Par principe, pour effec-

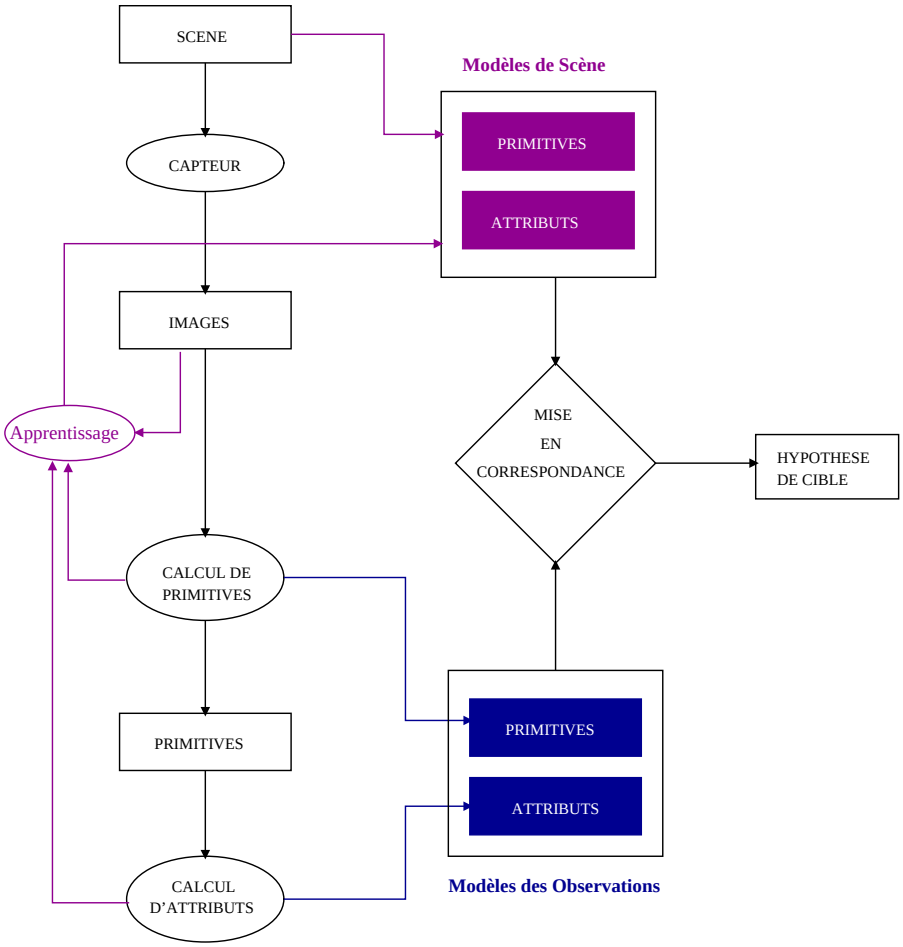


FIG. 2.1 – Schéma de principe de la fonction de détection, de reconnaissance et d'identification de cibles : l'appariement entre un modèle de scène et un modèle d'observation permet de formuler une hypothèse sur la cible. Les modèles de scène sont connus *a priori* ou extraits de l'image par apprentissage. Les modèles d'observation sont évalués à partir des données image sous forme de primitives ou d'attributs.

tuer une comparaison, il faut considérer des entités semblables. Aussi, pour réaliser la fonction DRI, il faut mettre en correspondance des modèles de même nature provenant d'espaces différents.

Toutefois des distinctions sont à considérer. Ces modèles diffèrent suivant la phase concernée (détection, reconnaissance ou identification) et donc suivant le niveau de détail disponible, mais également en fonction de la nature de la caractérisation. En effet, il est important de distinguer :

- la modélisation globale qui caractérise la globalité de l'objet. Elle est robuste au bruit mais présente de sérieuses limitations opérationnelles relatives aux phénomènes d'occultation des cibles.
- la modélisation locale qui s'intéresse à une partie de l'objet. Elle permet d'être plus robuste à des phénomènes d'occultation. En revanche, elle est beaucoup plus sensible au bruit.

Au travers d'une brève étude bibliographique, nous allons analyser les différentes solutions proposées pour caractériser une cible dans le contexte DRI, suivant le type d'informations manipulées, l'étape considérée et la nature de la modélisation.

2.2.1 Les modèles de scène

Un objet est défini par un ensemble de caractéristiques qui lui sont spécifiques. Celles-ci varient d'un objet à l'autre et d'un contexte à l'autre. Pour construire ces modèles, il existe deux possibilités représentées sur la figure 2.1 :

- soit, les modèles sont directement formalisés par l'expert, et dans ce cas ils se réfèrent à une réalité. L'espace d'exploration est la scène réelle.
- soit, les modèles sont construits par apprentissage. L'espace d'exploration est alors l'image.

La première approche ne prend pas en compte les imperfections d'acquisition liées au capteur. Les risques d'erreur sont alors très élevés. Cette modélisation donnée par exemple sous forme de modèles fil de fer est très proche de la réalité physique. Elle est donnée par l'expert en fonction des caractéristiques de l'objet.

Dans la modélisation par apprentissage, les conditions de perception sont intégrées, puisque le modèle est construit à partir des images issues du capteur. Ceci diminue les risques d'erreurs dans le processus de décision. Mais cette modélisation ne reflète qu'une vue partielle et imparfaite de la réalité.

Ces deux approches apparaissent d'ores et déjà assez complémentaires dans la prise en compte d'une réalité physique de l'environnement et des imperfections du système d'acquisition. Nous allons procéder maintenant à une analyse plus fine de ces deux approches. Nous distinguons différentes modélisations suivant le type d'in-

formations manipulées et leur chronologie d'apparition dans le cycle DRI.

Les modèles géométriques :

Ils traduisent une information sur la forme de l'objet.

Ils donnent une représentation de l'objet sous la forme d'une silhouette, d'un contour, d'un ensemble de surfaces ou de volumes si le système de perception fournit des informations tridimensionnelles. Ces indices perceptuels sont stockés sous forme d'arêtes, de modèles "fil de fer" pour les informations de type contour. Ils se déclinent sous forme de modèles à facettes [Stevens et Beveridge, 1997] ou de modèles CAD pour l'information surface [Olson et Huttenlocher, 1997; Der et Challappa, 1997]. Cette caractérisation sous forme de contours décrit l'objet dans sa globalité et pose le problème de la robustesse vis-à-vis des transformations géométriques, des changements d'échelle et des occultations. Elle nécessite, de plus, une base de données relativement importante puisque chaque objet doit être représenté suivant toutes les attitudes possibles.

Les modèles radiométriques :

Ils donnent une caractérisation de la surface "apparente" de l'objet

Les modèles radiométriques caractérisent la couleur de l'objet, sa luminance. Ils sont très variables au cours du temps. Ils sont notamment fonction de la luminosité et des phénomènes atmosphériques, aussi l'expert aura une grande difficulté à spécifier ce paramètre. L'apprentissage est donc une technique d'acquisition de connaissances plus appropriée. Stevens et Beveridge ont notamment proposé un apprentissage pour distinguer un pixel cible d'un pixel fond [Stevens et Beveridge, 1997], dans une phase de détection. Des travaux récents de M.R Stevens soulignent l'intérêt d'un imageur couleur pour la reconnaissance dans un contexte multicible [Stevens, 1999].

L'évolution vers les modèles physiques :

Ils offrent une caractérisation des grandeurs intrinsèques de l'objet

Les matériaux composant un objet ont des propriétés de texture et d'absorption de la lumière. Leur perception reste très sensible au choix du capteur et aux conditions d'acquisition des images.

Nandhakumar associe localement en chaque point d'intérêt un vecteur d'invariants thermophysiques [Nandhakumar *et al.*, 1997]. Chaque point caractérise une texture différente du véhicule telle que la carrosserie, les pneus, les vitres... Ces paramètres caractérisant partiellement l'objet permettent une robustesse à des variations de luminosité notamment.

Les modèles physiques peuvent également être donnés sous forme d'une caractérisation globale avec, par exemple la cinématique de l'objet. Le mouvement peut

être défini par une trajectoire 3D, ou des informations cinématiques [Castello *et al.*, 2000]. Il fournit des informations sur la présence d'une cible ou sur son type. Le modèle physique associé peut être spécifié soit par apprentissage, soit par l'expert directement.

Les modèles symboliques :

Les modèles symboliques viennent en complément des modèles numériques pour améliorer l'interprétation. Ils rassemblent des renseignements sur le contenu de la scène observé. Ils apparaissent dans les phases de reconnaissance et d'identification, soit sous forme de règles, soit sous forme de graphes.

- Les modèles structurels

Ils décrivent et spécifient les relations de composition d'un l'objet.

De niveau symbolique plus élevé que les précédentes modélisations, les modèles structurels donnent une description précise d'un objet en spécifiant les éléments qui le composent et les relations entre ces différents éléments, hypothèse faite que les différents sous-objets constituants puissent être isolés et détectés. Cette connaissance est souvent représentée sous forme de graphes.

L'avantage principal des graphes est de fournir une représentation de la connaissance très structurée et hiérarchisée. Ils sont utilisés pour définir les objets complexes ou des dépendances logiques entre leurs composantes. On parle parfois de *graphe d'aspect*. Un graphe d'aspect donne une représentation qualitative qui énumère les différentes apparences d'une cible suivant l'attitude considérée et son type [Hamilton et Kipp, 1993; Olson et Huttenlocher, 1997; Das et Bhanu, 1998]. Un exemple de représentation est donné sur la figure 2.2.

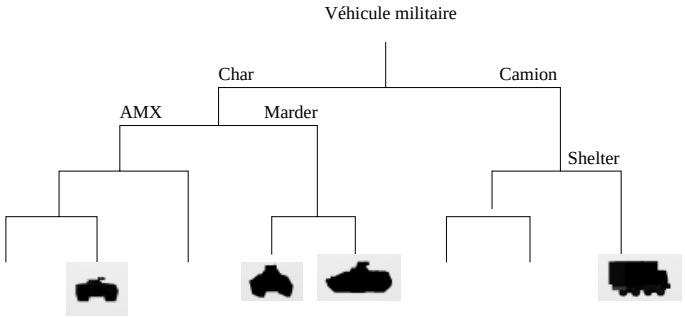


Fig. 2.2 – Description hiérarchique de véhicules militaires.

Il peut être obtenu de la manière suivante. On place la cible au centre d'une sphère imaginaire et on déplace le capteur sur la surface de celle-ci, tout en pointant son centre. Les régions de la sphère qui donnent la même apparence de la cible sont les noeuds du graphe. Les relations topologiques entre les régions sont décrites par

les arcs.

Une autre représentation possible est le *graphe de composition* des éléments de la scène [Acheroy et Mees, 1991]. Un noeud est un motif et un arc décrit une relation topologique entre deux motifs. J. Verly [Verly et Delanoy, 1996] donne une description hiérarchique des véhicules militaires. Entre chaque noeud du graphe, il y a soit une contrainte telle que "sur", "à côté de", "proportionnel à", soit un lien du type "spécialisation" ou "partie de". Un tel graphe associe donc des relations de spécialisation, des relations de composition et des relations spatiales.

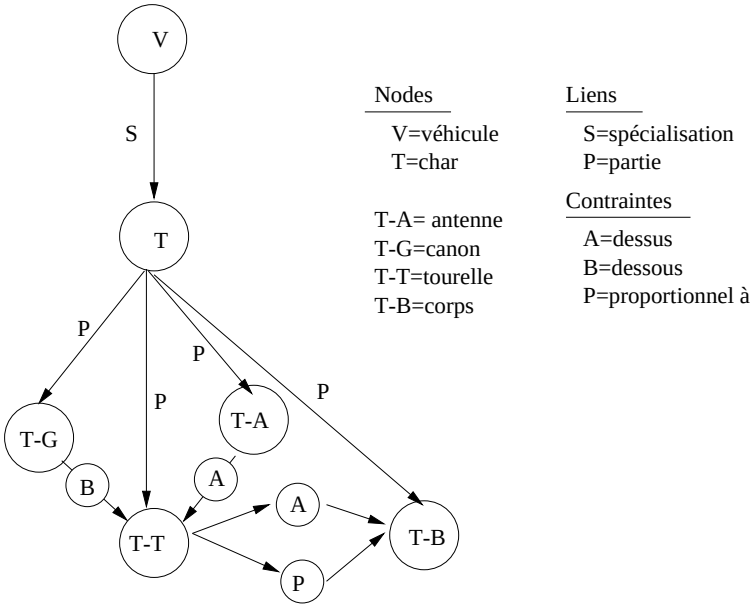


FIG. 2.3 – Graphe de composition et de spécialisation pour la reconnaissance de véhicules militaires.

Les graphes peuvent être utilisés pour représenter les relations possibles entre les noeuds. Les méthodes de calcul associées sont traitées dans la théorie des graphes. Les graphes constituent également un outil adapté à la représentation des relations de causalité et des relations temporelles. Il est possible d'associer aux arcs d'un graphe des conditions de transition, des notions de durée, de probabilité de transition, ou de degré de croyance pour représenter l'incertitude d'une relation de causalité.

- Les modèles contextuels

Ils spécifient les positions relatives des éléments de scène

Ils permettent de considérer la pertinence de la localisation d'un objet par rapport à son environnement [Bhanu, 1986]. La présence d'une voiture est fréquente sur les routes. Elle est beaucoup plus rare dans les champs, dans le ciel, ou sur la mer. Là encore cette connaissance peut être spécifiée par l'expert.

Cette liste n'est absolument pas exhaustive, mais elle permet de prendre conscience de la grande diversité des caractérisations possibles pour un même objet dans un même contexte. Suivant le type de capteur utilisé, les conditions d'acquisition, la qualité et la quantité d'informations disponibles, le choix des modèles sera différent.

Conclusion

Nous avons décrit différents modèles de scène utilisés au cours du processus de détection, de reconnaissance et d'identification. Le but est de trouver des critères qui permettent la meilleure caractérisation possible qui soit invariante par rotation, par translation, et changement d'échelle, par occultation... À l'heure actuelle, la solution miracle n'existe pas et les différentes modélisations mises en oeuvre répondent à certaines difficultés sans pour autant être robustes à toutes les situations. Les modèles proposés se distinguent suivant :

- les informations manipulées : géométriques, radiométriques, physiques, symboliques.
- la nature de la caractérisation : locale ou globale
- la construction du modèle : apprentissage ou spécification par un opérateur

Force est de constater l'importance des diversités et des compromis. Dans un système DRI, on va chercher en permanence à utiliser des informations complémentaires pour caractériser une cible. On aura tout intérêt à utiliser des modélisations manipulant des informations de natures différentes pour enrichir la connaissance. D'un autre côté, il faut établir un compromis entre les caractérisations locale et globale qui autorisent respectivement une robustesse au bruit et aux occultations éventuelles.

Nous allons poursuivre notre analyse, en étudiant les modèles d'observation. Nous garderons la même philosophie de description en considérant notamment la nature des informations manipulées.

2.2.2 Les modèles d'observations

Du point de vue de la modélisation des observations, un système DRI est représenté comme un système hiérarchique à trois niveaux d'abstraction.

- les attributs bas niveau représentent les primitives de l'image telles que les contours, les régions, et le mouvement. Ces attributs donnent une description très ponctuelle de la cible et sont fréquemment utilisés en phase de détection.

- une caractérisation intermédiaire permet le passage entre les informations bas niveau et les informations symboliques de haut niveau. Cette caractérisation est basée sur des notions diverses telles que la géométrie, la radiométrie ou encore la topologie. Elle permet de regrouper, de filtrer des primitives de l'image,

de fusionner des sources d'informations... Ces caractéristiques sont les éléments du processus de reconnaissance.

– les attributs haut niveau donnent une description globale et spécifique de chaque objet, et autorisent une identification de la cible.

La définition de ces différents niveaux d'abstraction ne signifie pas pour autant leur indépendance. Par exemple, le niveau intermédiaire peut fournir au bas niveau des informations contextuelles lui permettant de contraindre effectivement son traitement.

La modélisation des observations permet de construire une représentation symbolique à partir des mesures issues d'un ou plusieurs capteurs. L'espace d'exploration est à l'origine l'espace image. Le but est de trouver un ensemble de primitives ou d'attribut permettant la meilleure discrimination entre la cible et son environnement.

Cette modélisation doit être cohérente avec la modélisation de l'objet donnée par l'expert ou obtenue par apprentissage (modèle de scène). Aussi, on retrouvera les mêmes modélisations suivant les informations manipulées à savoir les modélisations géométriques, radiométriques, physiques, structurelles et contextuelles. On insistera également sur la distinction entre les approches locales et globales.

Nous allons maintenant détailler ces différentes approches. Pour chaque modèle, nous spécifierons les références bibliographiques, le type de caractérisation (locale ou globale), la phase dans laquelle il est utilisé (détection (D) ou reconnaissance et identification (R-I)), mais également les avantages et les inconvénients qu'il présente.

Les modèles géométriques

Ils ont la propriété de caractériser dans l'espace image
la forme d'une région ou d'un contour.

Les modèles géométriques sont très souvent utilisés dans la fonction DRI. Généralement, dans le processus classique d'analyse d'images, ces modélisations sont issues des mesures consécutives à une segmentation et au calcul d'attributs ou de vecteurs d'attributs pour tout objet détecté. On distingue les approches basées sur les primitives contours et sur les primitives régions (cf. figure 2.4).

♦ Les approches basées sur des primitives de type contour :

Elles nécessitent des connaissances suffisantes sur la cible, avec notamment la présence de contours significatifs. Elle ne peut être envisagée pour la détection de petites cibles. Chaque fois que l'on parlera de détection avec des primitives de type contour, implicitement cela se référera à des véhicules représentés en général par plusieurs centaines de pixels.

C. Olson et D. Huttenlocher préconisent de modéliser les objets par leur contour, sachant qu'une composante liée à la direction du gradient est associée à chaque pixel de l'image [Olson et Huttenlocher, 1997]. M.R. Stevens et J.R. Beveridge proposent également de caractériser les objets par leurs silhouettes [Stevens et Beveridge, 1997]. Cette information est ensuite complétée par des attributs statistiques pour diminuer le nombre de fausses alarmes.

Rao caractérise une cible par une liste chaînée de points anguleux issus de son contour [Rao, 1993]. À chaque contour est associée une mesure de saillance basée sur la longueur, la régularité et le contraste. Seuls les contours les plus significatifs sont validés et participent à la détection. Une cible peut également être caractérisée par un ensemble de sommets ou de points anguleux. C'est précisément l'approche préconisée par [Young *et al.*, 1998]. À chaque sommet est associé un vecteur de données géométriques.

M. Leung et T. Huang exploitent également les contours des roues des objets pour reconnaître les cibles [Leung et Huang, 1991]. À chaque objet est associé un vecteur de données précisant le nombre de roues, leur taille et leur position. Cette caractérisation qui est globale, reste très sensible aux occultations et aux changements d'échelle.

Afin d'augmenter la robustesse aux occultations Gupta propose une caractérisation locale des contours de l'objet [Gupta *et al.*, 1995]. Un objet est représenté par un ensemble de vecteurs d'attribut délimité par les points de forte courbure. Les composantes du vecteur sont fonction de sa longueur et de la distance aux extrémités. Cette méthode qui utilise une caractérisation locale pour reconnaître les cibles reste sensible aux changements d'échelle.

Une autre méthode consiste à décrire entièrement le contour par un ensemble de coefficients complexes tels que les descripteurs de Fourier. Cette description présente l'avantage d'être réversible (il est possible de reconstruire le contour à partir des descripteurs), de donner une représentation multi-échelle et d'être invariante par rotation, translation et homothétie du plan après normalisation. Cette méthode a été largement utilisée en reconnaissance de cibles [Das et Bhanu, 1998; Henocq et Burel, 1993] et plus spécifiquement pour la reconnaissance et l'identification d'avions.

♦ Les approches basées sur des primitives de type région

Cette approche peut aussi bien être utilisée en phase de détection avec de petites cibles, qu'en phase de reconnaissance ou d'identification. Elle apparaît donc déjà moins contraignante que l'approche de type contour.

Drake propose une approche basée sur le calcul de paramètres géométriques à

partir de l'image segmentée. Les mesures telles que la surface (S), le périmètre (P), la surface de la boîte englobante (SB), ainsi que leurs rapports d'allongement et de compacité (S/SB, P/S, P^2/S) permettent une caractérisation relativement simple des objets. Cette méthode a été utilisée pour reconnaître des véhicules blindés dans un contexte multi-capteur [Drake et Kim, 1995].

Il est également possible de caractériser la géométrie en travaillant directement sur les niveaux de gris dans l'image. Par exemple, le calcul de la signature fractale permet de quantifier l'irrégularité d'une surface. Dans le cadre de la détection automatique de cible, Espinal utilise une mesure basée sur la décomposition en ondelettes de Gabor [Espinal *et al.*, 1998].

Un récapitulatif des approches proposées est présenté dans le tableau Fig. 2.4.

Les modèles statistiques

Ils donnent la répartition des niveaux dans l'espace image ou dans un espace transformé (fréquentiel).

Différentes modélisations statistiques sont utilisées. Une première méthode consiste à utiliser les matrices de co-occurrence pour différencier des textures entre elles. Cette représentation contient les moyennes d'espace du second ordre. Auborn *et al.* proposent de calculer l'histogramme de la matrice de co-occurrence pour détecter les cibles potentielles dans l'image [Auborn *et al.*, 1993]. Ce procédé a l'avantage d'être simple et implémenté en temps réel. En revanche, il pose des difficultés de robustesse vis-à-vis du bruit et de la richesse du contexte.

En statistique, les moments permettent de caractériser la distribution d'une variable aléatoire, alors qu'en reconnaissance des formes, ils représentent la distribution spatiale de la forme. Cette modélisation se situe donc à la limite entre une caractérisation statistique et géométrique. Prokop *et al.* ont proposé un état de l'art concernant les différentes techniques utilisées pour la reconnaissance des objets non occultés [Prokop et Reeves, 1992]. Ces méthodes sont en général invariantes par transformations géométriques, mais restent très sensibles au bruit, et aux occultations.

Il existe également des modèles statistiques basés sur une approche locale et notamment sur le calcul de vecteurs d'attributs. Chaque pixel de l'image est alors caractérisé par un vecteur représentant le résultat d'un filtrage directionnel (probe based approach) [Der et Challappa, 1997]. Cette information associée à une information géométrique de type contour permet de déterminer un rapport de vraisemblance entre l'hypothèse cible et l'hypothèse fond. Cette méthode montre l'importance de la complémentarité entre informations de natures différentes.

La caractérisation par des attributs de couleur peut également apporter une information supplémentaire. Scribner *et al.* proposent d'utiliser plusieurs bandes

Modèles	L/G	D - R - I
Contours		
- [Olson et Huttenlocher, 1997]	– Global	– Reconnaissance : une mesure d'orientation est associée à chaque pixel du contour (distance de Hausdorff).
- [Stevens et Beveridge, 1997]	– Global	– Reconnaissance : l'information contour est utilisée pour décrire la forme. Cette caractérisation est enrichie par des données statistiques pour reconnaitre les cibles.
Régions		
- [Espinal *et al.*, 1998]	– Global	– La détection se fait par calcul de la signature fractale. Cette méthode basée sur les ondelettes permet de caractériser l'irrégularité d'une surface.
- [Drake et Kim, 1995]	– Global	– La détection est réalisée par un algorithme basé sur les invariants statistiques et les moments. La reconnaissance exploite les grandeurs géométriques liées à la région telles que la longueur, la largeur, le périmètre, la surface, la moyenne, l'écart type ...
Vecteurs d'attributs		
- [Rao, 1993]	– Lobal	– Détection : extraction des contours, sélection des plus significatifs. Le critère de détection est basé sur la mise en correspondance des coins 2D avec les sommets 3D.
- [Leung et Huang, 1991]	– Global	– Détection : extraction des contours et recherche des formes circulaires dans des zones d'intérêt. Reconnaissance : mise en correspondance à partir d'un vecteur d'attributs composé de trois types de paramètres : le nombre de roues, leur taille, leur position.
- [Gupta *et al.*, 1995]	– Lobal	– Détection : extraction des contours. Reconnaissance : mise en correspondance de vecteurs composés de distances euclidiennes.
- [Das et Bhanu, 1998], - [Henocq et Burel, 1993]	– Global	– Reconnaissance : description du contour par les coefficients de Fourier

FIG. 2.4 – Récapitulatifs des modélisations géométriques proposées pour réaliser le processus DRI.

spectrales pour rechercher les cibles dans l'image et de fusionner les données ainsi obtenues [Scribner *et al.*, 2000]. On insiste ici sur la complémentarité des informations multi-bandes.

Il est également possible de représenter la répartition des niveaux dans l'espace fréquentiel. Cette approche est couramment proposée dans la littérature. La décomposition en ondelettes de Gabor est la technique la plus souvent utilisée [Wu et Bhanu, 1997; Jain *et al.*, 1997; Hecht-Nielsen *et al.*, 1994]. La mesure obtenue est locale en fréquence et directionnelle en espace, ce qui impose de faibles transformations géométriques de la cible. Afin de résoudre ce problème, Wu et Bhanu proposent d'utiliser un filtre multidirectionnel et multi-échelle [Wu et Bhanu, 1997]. Cette technique suppose une discrétisation de l'espace. En revanche, la décomposition en sous-bandes permet d'avoir accès aux informations hautes et basses fréquences, tout en conservant l'information spatiale. [Jain *et al.*, 1997] proposent également d'associer une mesure d'énergie basée sur les coefficients de Gabor pour modéliser les objets.

Les modèles physiques

Le modèle physique a la particularité de mesurer une grandeur intrinsèque à l'objet.

La cinématique de l'objet est une grandeur qui peut permettre de distinguer les cibles entre elles. Par exemple, pour une vue de profil, un avion en vol ira toujours plus vite qu'un char qui se déplace. Castello *et al.* utilisent un algorithme basé sur le flot optique (Regularized CDWT optical flow) pour détecter les petites cibles en mouvement dans une séquence d'images [Castello *et al.*, 2000]. Cette technique est contrainte à de petits déplacements. Une seconde approche possible est d'utiliser le mouvement pour se focaliser dans l'image.

Nandhakumar utilise les propriétés thermophysiques des matériaux pour reconnaître les objets [Nandhakumar *et al.*, 1997]. Le principe de conservation de l'énergie entre la surface et l'environnement est considéré. Il permet de définir en chaque point caractéristique un tenseur à cinq composantes. Cette méthode montre une certaine robustesse vis-à-vis des variations d'intensité. En revanche le problème de la sélection optimale de l'ensemble des points caractéristiques reste une difficulté.

2.2.3 Conclusion

Force est de constater une très grande diversité de méthodes pour réaliser la fonction DRI. On distingue les modèles de scènes et les modèles d'observation. Le modèle de scène, défini *a priori* ou par apprentissage, donne une représentation de la cible dans le contexte d'analyse. Le modèle d'observation est le résultat des mesures effectuées dans l'espace image sous forme de primitives ou d'attributs.

Les modèles se distinguent par la nature des informations manipulées. Qu'elles soient géométriques, statistiques, physiques ou encore symboliques, ces informations caractérisent partiellement la cible. Souvent, il s'avère nécessaire d'utiliser plusieurs méthodes pour détecter, reconnaître ou identifier une cible. Ceci se traduit par une complexité plus grande.

Dickinson, Pentland et Rosenfeld ont étudié le compromis à établir entre [Dickinson *et al.*, 1992] :
- la complexité des primitives d'observation,
- la complexité du modèle de scène,
- la complexité de la recherche,
- la flexibilité du modèle,
- la confiance dans la reconnaissance,
- la difficulté de l'extraction des primitives.

Plus les primitives sont complexes, et plus le modèle est simple. De plus la recherche est simplifiée puisque le nombre de primitives à rechercher est plus faible. En revanche, le système devient moins fiable puisque l'extraction de telles primitives complexes reste peu robuste.

Un compromis supplémentaire est à faire, concernant les modélisations locale et globale. Cet aspect est très important. En effet, la caractérisation globale a de sérieuses limitations opérationnelles relatives aux phénomènes d'occultation des cibles. D'un autre côté, si la caractérisation locale de la cible va permettre d'être plus robuste à ces phénomènes d'occultation, elle sera en revanche beaucoup plus sensible au bruit. Par conséquent, il apparaît d'ores et déjà qu'un compromis entre ces modèles globaux et locaux est inévitable.

Pour construire notre système, nous avons retenu l'ensemble de ces remarques. Nous avons utilisé des informations de natures différentes quelle que soit l'étape considérée, prenant en compte à la fois une caractérisation locale et globale. Toutefois les contraintes liées au domaine d'application ont parfois imposé des choix. En phase de détection, la taille des cibles de l'ordre de quelques dizaines de pixels a notamment limité l'utilisation d'une approche de type contour. C'est pourquoi nous avons plutôt basé notre raisonnement sur une approche de type région pour la modélisation des cibles.

D'autres problèmes se posent dès lors que l'on considère de petites cibles dans l'image. En effet, celles-ci se trouvent noyées au sein d'une multitude d'informations et le risque d'erreur de détection est élevé. Il s'avère dans ce cas judicieux de restreindre l'analyse par l'emploi de connaissances contextuelles permettant de focaliser l'attention sur telle ou telle partie de l'objet. Pour chercher une voiture, par exemple, il peut être intéressant de chercher d'abord des routes (éléments structurels forts dans l'image) et de focaliser ensuite l'attention sur les supports de ces routes.

Enfin, il y existe un dernier point qu'il est important de mentionner : l'adaptation. Il convient de distinguer deux formes complémentaires d'adaptation qui concernent respectivement les méthodes de détection et la mise à jour des connaissances. Il s'agit dans le premier cas de tenir compte du fait qu'un objet peut se trouver placé dans un contexte non prédictible *a priori* (un véhicule sur une route, dans un champ...). Il s'agit dans le deuxième cas de considérer les diversités d'apparences possibles pour un même objet dues aux conditions d'acquisitions (cible vue de face ou de profil...) et aux perturbations éventuelles (variations de luminostité par exemple) .

Cette étude bibliographique a mis en évidence un grand nombre de modélisations possibles pour les cibles. Ceci se traduit également par une grande diversité dans les méthodes de mise en correspondance sur lesquelles nous allons désormais porter notre attention.

2.3 Les techniques de mise en correspondance

Un très bon état de l'art a été proposé par Z. Zhang concernant le problème de la mise en correspondance dans les problèmes de vision par ordinateur et spécifiquement pour la reconnaissance d'objets [Zhang, 1993]. Le problème posé revient à calculer une distance entre le modèle de l'objet (issu du modèle de scène) et le modèle d'observation. Nous présentons dans ce paragraphe les principales méthodes utilisées dans la mise en œuvre de la fonction détection, reconnaissance et identification de cibles [Li *et al.*, 1998]. Nous distinguons cinq approches différentes.

2.3.1 La théorie des graphes

La littérature concernant les arbres de décision est très riche [Murthy, 1998; Safavin et Landgrebe, 1991; Dattatreya et Kanal, 1985]. Pour effectuer la mise en correspondance, plusieurs approches sont possibles :

• Recherche de sous graphes
L'appariement de graphes est une technique utilisée pour reconnaître les objets dans une scène. Un premier graphe représente le modèle d'objet. À chaque primitive est associé un noeud et chaque arc décrit une contrainte entre deux primitives. Le second graphe représente les mêmes informations mais relativement aux informations issues des données capteur. La mise en correspondance consiste alors à trouver des sous-graphes isomorphes dans les deux graphes.

• Construction de graphes de compatibilité
Cette méthode consiste à construire un graphe d'appariement [Pampagnin, 1990]. Les noeuds représentent les appariements capteur/primitives et les arcs représentent la compatibilité entre deux couples. Deux appariements sont compatibles si les relations existantes entre deux primitives issues du modèle sont également validées pour deux primitives sensorielles.

H. Moissinac-Massenat propose en état de l'art des méthodes de mise en correspondance et de comparaison de graphes [Moissinac-Massenat, 1996] dans le cadre de l'interprétation de scènes urbaines et de la fusion de données. Les graphes permettent la représentation simultanée de l'information symbolique (le graphe) et numérique (mesure de confiance attachée à chaque noeud).

2.3.2 Les techniques de seuillage

Le seuillage

Ce procédé revient à fixer une distance maximum entre un modèle et une observation. D. Casasent et A. Ye [Casasent et Ye, 1997] ont développé différentes techniques de seuillage pour détecter des cibles. Ils proposent un algorithme qui fusionne les informations ("ET") et qui utilise un seuillage adaptatif pour sélectionner les cibles. Des variantes sont également proposées. Celles-ci sont basées sur une fusion issue de la logique floue et une recherche de pics d'intérêt (peak sorting). Elles ont permis notamment de diminuer le nombre de fausses alarmes.

Le hachage géométrique

Le principe de cette technique repose sur la construction de tables de hachage. Lamdan et Wolfson [Lamdan et Wolfson, 1988] proposent d'identifier un objet à partir de points d'intérêt caractérisés par des attributs invariants par rotation, translation et changement d'échelle. Les entrées de la table de hachage sont des bases géométriques dans lesquelles sont exprimés ces invariants [Tsai, 1996]. I. Rigoutsos et R. Hummel proposent une formulation bayésienne du hachage géométrique qui leur permet d'exprimer le vote en fonction d'une estimation des erreurs de positionnements des différents points [Rigoutsos et Hummel, 1995]. L'inconvénient majeur de cette approche reste l'occupation de l'espace mémoire.

2.3.3 Les calculs de distance

De nombreuses mesures sont proposées dans la littérature. On prendra quelques exemples pour illustrer cette approche.

La mesure de Chamfer

Elle minimise la somme des distances entre les pixels contour du modèle et les pixels contour de l'observation [Barrow *et al.*, 1977; Borgefors, 1988]. Cette technique est étroitement liée à la minimisation de la mesure de Hausdorff généralisée, ou encore à la minimisation des K plus grandes distances. La mesure de Chamfer somme les distances sur tous les pixels. Notons qu'elle n'est pas robuste aux occultations.

Les scores de corrélation

J.E. Auborn propose d'utiliser une technique de corrélation pour mesurer la similarité entre deux objets [Auborn *et al.*, 1993]. Des scores de corrélation sont établis à partir des calculs de distance euclidienne ou de produits croisés entre les attributs issus respectivement des modèles de scènes et d'observation. Les applications de cette technique sont souvent liées à une modélisation par apparence qui utilise comme espace de représentation les espaces vectoriels.

La mesure de confiance

M.K. Leung et T.S. Huang calculent une mesure de confiance entre des attributs [Leung et Huang, 1991]. Chaque cible est caractérisée par un vecteur d'attributs A=Att et, pour chaque attribut, une mesure de confiance C est calculée. Sachant que l'indice "obs" référence le modèle d'observation et "sc" le modèle de scène, on a :

$$C = 1 - \frac{Att_{obs} - Att_{sc}}{Att_{sc}}$$

La confiance totale est alors égale au produit des différentes confiances obtenues pour chaque attribut.

Il est possible de changer d'espace pour calculer la distance. Cette technique consiste à définir un espace vectoriel, à partir des vecteurs propres, dans lequel les attributs sont les plus discriminants possible. Le procédé utilisé est par exemple l'analyse en composantes principales. Hsu *et al.* utilisent cette technique comme une première étape dans la définition d'un réseau de neurones [Hsu *et al.*, 1994].

2.3.4 Les classifieurs

Les techniques de classification sont nombreuses [Duda *et al.*, 2001]. Dans le domaine de la DRI , deux approches se distinguent. Elles sont basées sur les réseaux de neurones ou sur la classification bayésienne.

Les réseaux de neurones

Les réseaux de neurones sont des techniques très largement utilisées dans le domaine de la DRI [Henocq et Burel, 1993; Rogers *et al.*, 1995; Ernisse *et al.*, 1997; Wang *et al.*, 1998; Young *et al.*, 1998], pour fusionner les données numériques. Un état de l'art a été proposé par Roth [Roth, 1990].

• Ernisse utilise un réseau de neurones pour reconnaître des cibles [Ernisse *et al.*, 1997]. Ce réseau (Multi Layers Perceptron) prend en entrée les paramètres caractéristiques extraits de la région d'intérêt et propose une classification finale en sortie. Le réseau se construit à partir de bases d'apprentissage qui doivent être suffisamment riches et représentatives des situations potentielles.

• Wang propose une architecture basée sur la mise en œuvre de différents réseaux de neurones [Wang *et al.*, 1998]. Chaque réseau utilise des caractéristiques locales (approche hiérarchique) pour formuler une hypothèse. Chaque réponse est combinée pour donner la décision finale.

Classifieurs probabilistes

Nair et Aggarwal proposent une méthode de reconnaissance de véhicules basée sur une description hiérarchique des objets [Nair et Aggarwal, 2000]. La connaissance est structurée sous forme d'arbre alors que le processus de reconnaissance est basé sur la mise en œuvre de classifieurs bayésiens. Le schéma de principe est présenté sur la figure 2.5.

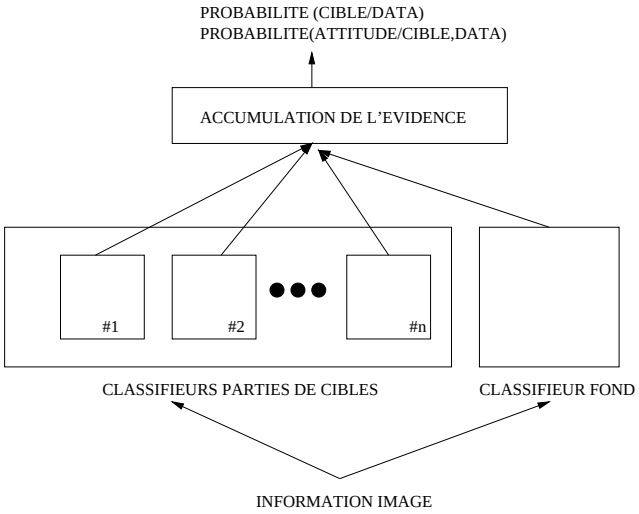


Fig. 2.5 – Reconnaissance par partie de la cible. Le classifieur est composé de différents modules, chacun représentant la fonction de probabilité conditionnelle d'une partie spécifique de l'objet

La reconnaissance est réalisée à différents niveaux hiérarchiques ; chaque niveau est constitué par un module "expert" qui est entraîné à reconnaître une partie d'un objet.

D. Buede et Girardi proposent une comparaison très intéressante entre les approches basées sur la théorie de l'évidence et les approches bayésiennes dans le cadre de la reconnaissance de cibles [Buede et Girardi, 1997]. Le formalisme bayésien nécessite la définition complète du modèle probabiliste mais le cadre d'étude est plus rigoureux et s'avère plus rapide. Nous reviendrons sur ces approches probabilistes dans le chapitre suivant.

2.3.5 Conclusion

Le problème de la mise en correspondance de modèles est un thème très largement abordé dans la littérature, tant par la diversité des connaissances utilisées, leur structuration, que par la variété des étapes et des niveaux d'interprétation où ces connaissances viennent se greffer. Nous avons distingué pour la fonction DRI quatre méthodes différentes : les graphes, les techniques de seuillage, les calculs de distances et enfin les classifieurs.

Concernant la structure de graphe, elle nécessite une mise en œuvre relativement complexe. Elle n'est pas adaptée à des situations où l'information présente est très faible comme c'est notre cas en phase de détection. En revanche, elle présente des avantages non négligeables pour la représentation des connaissances et la gestion de l'incertitude. Et c'est précisément pour cette raison que nous retiendrons cette approche. En effet les arbres peuvent être utilisés pour représenter des liens purement logiques, comme par exemple dans le cas d'arbres de probabilités. Ce type d'outil est très utile pour illustrer les dépendances entre variables et propager l'incertain.

Les techniques de seuillage permettent de mesurer l'écart entre une mesure et un seuil établit soit *a priori*, soit par apprentissage. Ces techniques restent en générale très sensible aux perturbations et nécessite par conséquent le choix de seuils adaptatifs ou d'attributs invariations par transformations.

Pour mesurer la distance entre le modèle de scène et le modèle d'observation, nous avons retenu l'approche par calcul de distances qui présente l'avantage d'être simple et facile à mettre en œuvre. Nous présenterons dans le paragraphe 4.6.2 la mesure utilisée.

Les classifieurs sont adaptés aux problèmes de la DRI. Toutefois, leur mise en œuvre reste complexe et repose sur l'apprentissage et le choix des pondérations toujours difficiles à maîtriser.

Parallèlement à la diversité des modèles, de nombreuses techniques ont été développées pour manipuler et combiner ces informations. L'ensemble de ces approches propose une classification, mais n'inclut pas le processus décisionnel que nous allons détailler maintenant.

2.4 Processus de reconnaissance

Jusqu'à présent, nous avons présenté des méthodes intervenant dans la phase d'appariement. Elles permettent de formuler des hypothèses à partir d'une mise en correspondance entre le modèle d'objet et le modèle d'observation. Avec le processus de reconnaissance, l'objectif est d'inclure un processus de décision pour gérer les hypothèses courantes.

La technique de reconnaissance peut être perçue comme un problème d'optimisation réalisé soit dans l'espace d'appariement, soit dans l'espace des positions de l'objet [Jonquières, 2000]. Quatre méthodes principales peuvent être dégagées : la relaxation, la recherche combinatoire, la recherche heuristique et le calcul d'une fonction coût/bénéfice. Nous allons préciser ces différentes approches.

2.4.1 La relaxation

Elle consiste à réaliser un filtrage itératif. La première étape est l'appariement entre les primitives d'observations et les primitives de scènes. Un degré de cohérence est caractérisé par un poids associé à chacune des étiquettes. Ces poids peuvent être probabilistes (relaxation stochastique) ou binaires (relaxation discrète). À chaque itération, les contraintes sont testées et les poids évoluent. La convergence est atteinte lorsque les poids ne varient plus dans le temps. Les solutions envisageables sont obtenues à la fin du processus itératif. C.F. Olson et D.P. Huttenlocher proposent une approche markovienne dans le cadre de la reconnaissance automatique de cibles [Olson et Huttenlocher, 1997].

2.4.2 la recherche arborescente

J. Verly utilise cette structure pour construire un système de reconnaissance automatique de cibles [Verly et Delanoy, 1996]. Le module de mise en correspondance dispose de trois entrées : un évènement qui correspond à une silhouette extraite des données image, une liste de primitives constituant la silhouette, et les modèles hiérarchiques des objets. Le but est de trouver l'objet dans le modèle hiérarchique qui permet la meilleure mise en correspondance vis-à-vis de la silhouette et des primitives trouvées. Le meilleur appariement est doté d'une mesure de confiance. Lors de la mise en correspondance, les propriétés de la théorie des graphes sont utilisés, et notamment le "pruning" qui permet une optimisation de l'espace de recherche.

2.4.3 La prédiction et vérification d'hypothèses

La boucle DRI peut être composée d'appariements successifs. Avec le premier appariement, une hypothèse est formulée (notamment concernant l'attitude et la classe de la cible), avec le second, elle est validée ou non. L'intérêt de cette technique est d'introduire une boucle de rétroaction sous forme de prédiction-vérification. Cette boucle se positionne de différentes façon dans le processus DRI. Deux exemples sont présentés sur les figures 2.6 et 2.7.

Les méthodes multicapteurs utilisent, en général, une boucle de prédiction vérification pour identifier les cibles. En effet, des hypothèses formulées dans une bande de fréquence peuvent être confirmées ou non dans une autre bande de fréquence. L'intérêt du multicapteur est indéniablement dans la richesse des informations prises en compte.

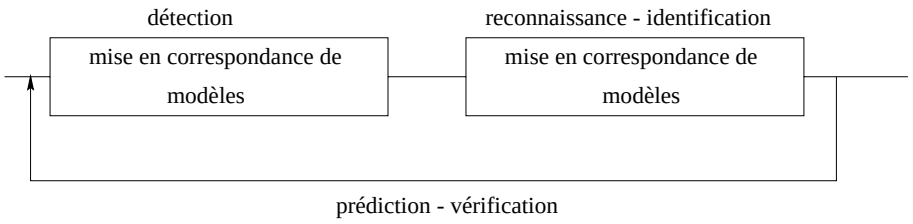


FIG. 2.6 – La boucle de rétroaction au sein de la fonction DRI inclut la phase de détection

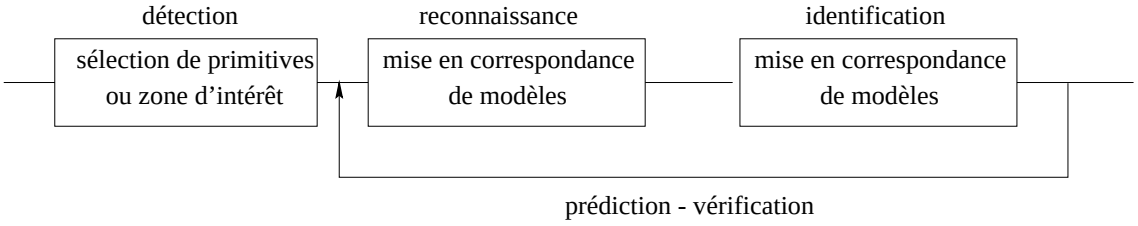


FIG. 2.7 – La boucle de rétroaction au sein de la fonction DRI n'inclut pas la phase de détection. Elle se fait au niveau de la reconnaissance et de l'identification

Dans la littérature propre à la DRI, de nombreuses approches utilisent ce procédé basé sur la génération et la prédiction d'hypothèses. Nous pouvons citer quelques exemples.

• Wu et Bhanu utilisent une représentation en ondelettes de Gabor pour réaliser la reconnaissance d'objets 3D [Wu et Bhanu, 1997]. Le processus de reconnaissance est itératif et se décompose en trois parties :

- positionnement d'une grille dans l'image, qui spécifie les points d'analyse. Pour cela, un algorithme de placement est mis en oeuvre de façon à minimiser la dissimilarité entre le modèle et l'objet. Cet algorithme permet d'estimer la position et l'échelle.
- mise en correspondance autorisant une flexibilité de la grille (modèle déformable).
- évaluation de la mise en correspondance.

Cette approche utilise deux boucles de contre-réaction. La position et l'échelle de la grille sont estimées dans l'algorithme de placement. Elles sont vérifiées lors de l'appariement. Le choix du modèle peut alors être remis en cause. De plus, lors des occultations éventuelles, la grille peut ne plus être adaptée. Dans ce cas, une transformation de cette grille est effectuée.

• J.R. Beveridge *et al.* ont développé un système de reconnaissance automatique de cibles qui fusionne les informations de profondeur (range image), les informations infrarouges et la couleur [Beveridge *et al.*, 1997]. Ce système utilise les informations

multicapteurs pour affiner les hypothèses concernant le type de cible et son attitude. Le schéma de principe est décrit sur la figure 2.8 :

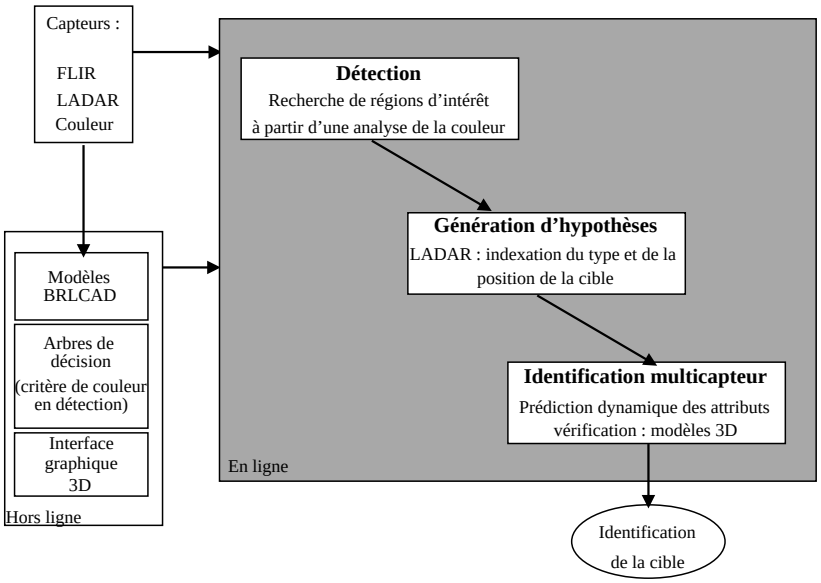


Fig. 2.8 – Schéma de principe d'un système multi-capteur de reconnaissance automatique de cibles développé par Beveridge *et al.*. Ce système se décompose en trois étapes distinctes : il réalise une détection, il génère une hypothèse sur le type et la position de la cible et finalement confirme son identité. Pour cela des données fournies *a priori* sont nécessaires telles que des modèles 3D, des arbres de décision (issus d'un apprentissage), et une interface graphique pour guider le module d'identification.

Le module d'identification est basé sur un principe de co-recalage. Il y a un recalage entre les données pour comparer les différentes sources spectrales. Mais il y a également un recalage entre le modèle 3D utilisé et la position du capteur. Le processus de reconnaissance est articulé ensuite suivant un schéma de prédiction-vérification.

• Drake et Kim proposent une stratégie de reconnaissance multispectrale basée sur des mécanismes de rétroaction [Drake et Kim, 1995]. La prédiction peut être étendue par des boucles de rétroaction qui autorisent la remise en cause de renseignements déjà collectés. Le schéma de principe est donné sur la figure 2.9.

Ce système de reconnaissance est composé de cinq modules différents. Le premier module est le système exécutif. Il coordonne les actions des différents modules de reconnaissance via des boucles de rétroaction. Cette boucle peut être locale, dans le cas de faibles performances chaque module peut être réactivé avec un nouveau jeu de paramètres. Elle peut également être globale et dans ce cas un module qui a déjà réalisé cette tâche, est activé.

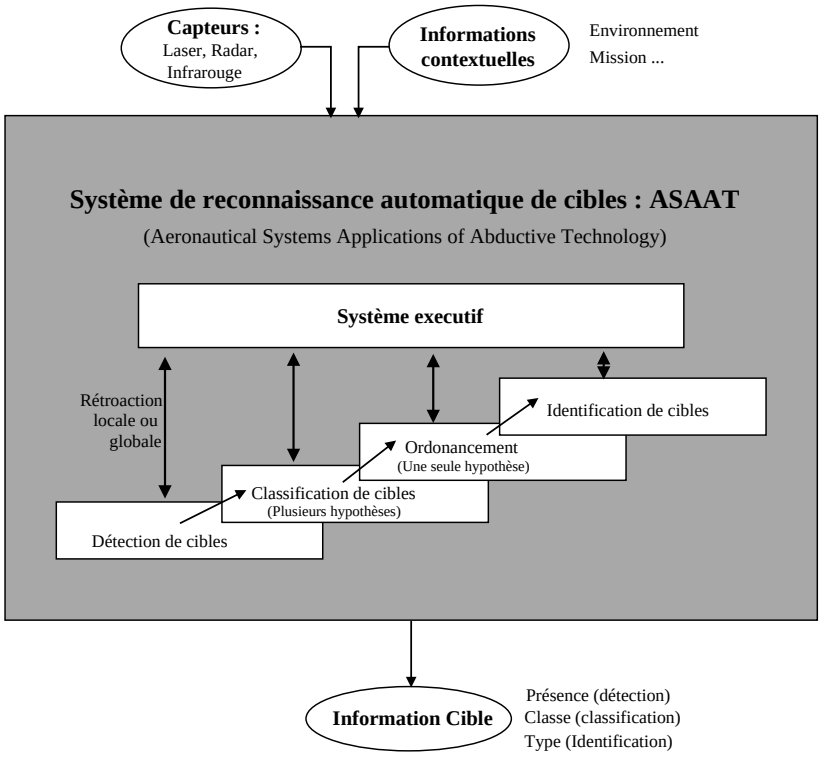


FIG. 2.9 – Le système de reconnaissance automatique de cibles : ASAAT.

Nous avons proposé quelques exemples pour illustrer l'intérêt des approches de type prédiction-vérification d'hypothèses. Elles permettent de remettre en cause des choix d'algorithmes, de paramètres, des résultats... Elles se présentent comme un moyen d'adapter le système à une situation et de remettre en cause des hypothèses.

2.4.4 La fonction coût/bénéfice

Roberts et Brown proposent une approche adaptative du problème de la DRI [Roberts et Brown, 1994]. Le but est de trouver une technique de configuration et d'apprentissage qui permette de sélectionner les algorithmes et les valeurs des paramètres pour réaliser le processus de DRI. Pour parvenir à la reconnaissance, il y a un questionnement sur l'utilité de tel ou tel algorithme.

L'architecture est basée sur la mise en œuvre d'un algorithme génétique qui produit et apprend de nouveaux enchaînements d'algorithmes pour réaliser le processus DRI. Le schéma de principe est présenté sur la figure 2.10.

Le module de perception sélective est capable de choisir l'enchaînement d'algorithmes à exécuter ainsi que l'endroit où il est nécessaire de l'appliquer dans l'image pour réaliser la fonction DRI ou une sous-tâche intermédiaire. La gestion de l'incertitude au cours du processus suit une approche bayésienne. L'information

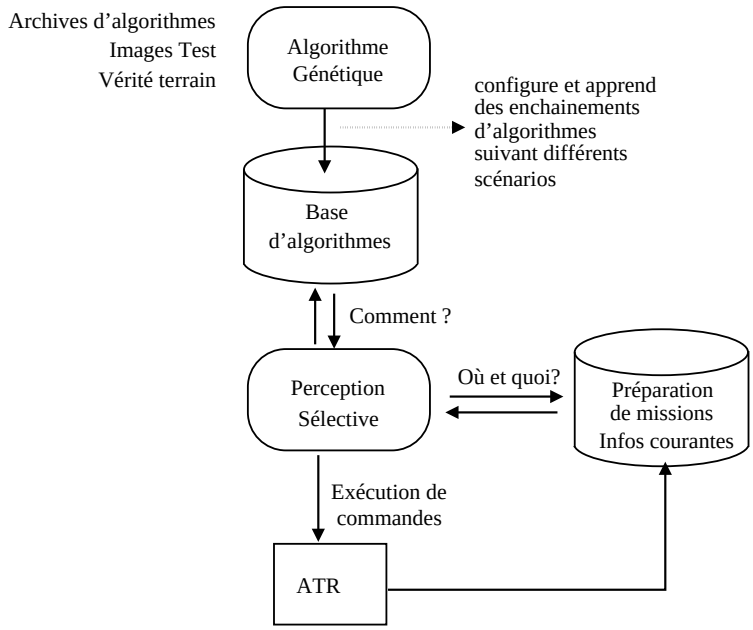


FIG. 2.10 – Diagramme représentant le système DRI proposé par Roberts et Brown.

a priori acquise durant la mission est représentée sous forme d'un réseau bayésien. Le choix de l'action à réaliser (Quoi ?), l'endroit où l'appliquer dans l'image (Où ?), et comment l'exécuter (Comment ?) sont déterminés par l'analyse d'une fonction coût/bénéfice.

Cette approche permet d'aller un peu plus loin dans la notion d'adaptation. Elle ne concerne plus uniquement le choix des méthodes et des paramètres, mais elle intègre la notion de but et de lieu. Cette fois on se pose la question de savoir quelle est l'information significative et où il faut la chercher dans l'image.

Le deuxième point important qui a été abordé dans ce système est l'utilisation des réseaux bayésiens pour à la fois structurer l'information et propager l'incertitude afin d'évaluer la fonction coût/bénéfice. Dans cette approche de fonction coût/bénéfice, on va plus loin que dans la prédiction et la vérification d'hypothèses. Car au lieu de remettre en cause une action on se pose la question de sa pertinence au préalable.

2.4.5 Conclusion

Nous avons présenté un ensemble de techniques réalisant un processus de reconnaissance automatique de cibles, incluant la gestion d'hypothèses et la prise de décision. L'approche sous forme de prédiction-vérification et de rétroaction est intéressante dans le sens où elle permet de remettre en cause des hypothèses, des algorithmes ou des paramètres. Elle autorise une adaptation en fonction de la situation courante. L'algorithme proposé par Roberts et Brown complète cette approche et notamment en ce qui concerne la notion d'adaptativité. De plus, ils donnent dans

un même schéma de représentation que sont les réseaux bayésiens, la structuration des connaissances et la propagation des incertitudes. Avec ces réseaux, on se pose la question de l'utilité d'une action connaissant la situation courante et les incertitudes présentes.

Les trois systèmes présentés suivent une démarche incrémentale dans la résolution du problème de vision. Ils procèdent progressivement par accumulation d'informations, vérification d'hypothèses et d'adaptation. Nous retiendrons ce principe d'incrémentalité dans la conception de notre système.

2.5 Conclusion

Après ce tour d'horizon présentant les diverses méthodes existantes pour modéliser les objets, pour effectuer une mise en correspondance et pour finalement prendre une décision, il reste un certain nombre de conclusions à tirer.

On distingue différentes modélisations. Elles diffèrent suivant les informations manipulées (géométrique, statistique ou physique), suivant leur type (modèle d'observations ou modèle de scène), mais également suivant leur nature (modèle local ou global). Les approches de type géométrique sont sans aucun doute les plus utilisées dans la littérature. Elles sont souvent associées à des informations radiométriques. La complémentarité des informations s'avère importante.

Comme le précise justement H. Moissinac-Massenat [Moissinac-Massenat, 1996], la combinaison des informations se fait très souvent en deux temps. Il y a une combinaison des informations numériques, suivie d'une utilisation des données symboliques. Et suivant la même philosophie, nous avons voulu incorporer ces connaissances dites de haut niveau très tôt dans le processus DRI. Il sera notamment intéressant de l'utiliser dès la phase de détection pour diriger l'analyse.

Les processus de reconnaissance sont bien souvent difficiles à cataloguer car ils intègrent différents procédés portant sur des hypothèses de natures très variables. Ces hypothèses concernent le type de la cible, son attitude, le choix ou encore les paramètres de l'algorithme.

L'approche la plus utilisée reste la prédiction et la vérification d'hypothèses. L'inconvénient majeur est souvent la gestion d'un grand nombre d'hypothèses, qui impose en grand nombre de combinaisons à examiner. Ce principe d'exploration associé à la rétroaction à l'avantage d'effectuer des remises en causes d'hypothèses, de choix d'algorithmes ou de paramètres. En cas de mauvaises performances, il propose de changer de stratégie plutôt que de poursuivre l'analyse avec des hypothèses erronées. Il permet une adaptation du système en fonction de la situation courante. Le calcul d'une fonction coût/bénéfice permet de compléter ce schéma adaptatif en se posant la question de l'utilité d'une action. Cette approche est particulièrement intéressante et nous la développerons dans la suite de notre manuscrit.

Nous avons vu en outre le besoin d'une analyse adaptée au contexte global, local, spatial pour répondre aux contraintes de la DRI [Bhanu et Jones, 1992; Dudgeon et Lacoss, 1993; Ratches *et al.*, 1997]. Seule une approche distribuée permet de réaliser cela. Elle permet de distribuer la connaissance et d'effectuer des traitements en parallèle et en concurrence. Aujourd'hui, le but n'est plus de développer de nouveaux algorithmes, mais davantage de les combiner pour tirer profit au maximum des informations qu'ils fournissent. La problématique est donc orientée vers le contrôle.

Nous avons évoqué aux cours de ce chapitre des questions relatives à la focalisation, la distribution, l'adaptation, l'incrémentalité et la gestion des incertitudes. Nous proposons désormais d'approfondir cette analyse, dans un contexte plus large regroupant les systèmes de vision.

Chapitre 3

L'approche proposée

Sommaire

3.1 Introduction 53
3.2 Vers un système de vision 53
3.3 Focalisation et adaptation 57
3.4 Distribution et coopération 59
3.5 L'incertitude 62
3.6 Incrémentalité et apprentissage 68
3.7 Conclusion 69

Résumé

Dans ce chapitre, nous proposons une approche décentralisée pour résoudre le problème de la détection, de la reconnaissance et de l'identification de cibles. Celle-ci s'est inspirée de plusieurs systèmes de vision existants et s'appuie sur de grands principes tels que la focalisation, l'adaptation, l'incrémentalité, la distribution, la coopération et la gestion des incertitudes.

L'approche est basée sur une architecture multi-agent qui pose le problème de la vision comme un problème d'ancrage dans un espace à trois dimensions : l'espace image, l'espace des buts et l'espace des méthodes. L'objectif du système est d'accumuler des informations pertinentes pour réaliser un but. Les informations collectées sont imparfaites et incomplètes. Aussi, la gestion des incertitudes est nécessaire et se fera à l'aide de réseaux bayésiens.

L'objectif de ce chapitre est de justifier nos choix par rapport à la littérature et de positionner notre travail dans le domaine des systèmes de vision.

3.1 Introduction

Les systèmes de vision ont pour vocation, à partir de données image hétérogènes, de permettre la compréhension de la scène observée. Cette scène est souvent caractérisée par un environnement non homogène, très variable et dont les conditions d'observation peuvent être quelconque. Bien évidemment cette généralité impose une complexité importante au sein des systèmes incluant des exigences d'évolutivité, des capacités à fournir des résultats intermédiaires, l'intégration de connaissances incertaines, ou encore la possibilité de s'adapter par le choix de stratégies, d'opérateurs, et de paramètres.

Progressivement les systèmes de vision ont quitté le schéma séquentiel d'interprétation construit autour de l'extraction de paramètres, de leur description symbolique et de l'interprétation sémantique, pour s'orienter vers des principes de contrôle [Draper *et al.*, 1996; Crevier et Lepage, 1997]. Ils ont alors été dotés de capacités à organiser leurs propres ressources en fonction d'une part des données observées et d'autre part de l'objectif à atteindre [Serfaty-Dutron et Blanc-Talon, 1999].

Dans ce chapitre, nous proposons une approche pour résoudre le problème de la détection, la reconnaissance et l'identification de cibles. Celle-ci s'appuie sur de grands principes tels que l'adaptativité, l'incrémentalité, la distribution, la coopération et la focalisation. Ces principes ont été introduits par C. Garbay [Garbay, 2000]. Nous proposons une justification de ces choix et présentons les approches dont nous nous sommes inspirées. L'objectif de ce chapitre est de positionner notre travail dans le domaine des systèmes de vision, et non plus uniquement dans le cadre strict que nous avons dressé pour la DRI.

3.2 Vers un système de vision

Comme le précise très justement V. Serfaty le mot contrôle a son revers. Ce mot *"... laisse penser qu'une démarche automatique a des chances d'aboutir, ce que la pratique infirme formellement aujourd'hui. Pour cette raison, il semble préférable d'utiliser la dénomination stratégie de vision..."*[Serfaty-Dutron et Blanc-Talon, 1999].

Pour suivre cette philosophie, nous proposons un système qui s'oppose à une démarche automatique mais en revanche qui procède de manière exploratoire. Le problème de vision est posé comme un problème incrémental de collecte d'informations au cours duquel on pose à chaque étape la question du où, quoi, comment :

- Où se situe l'information significative dans l'image ?
- Quelle est l'information significative ?
- Comment extraire cette information de l'image ?

En permanence, il faut poser le système de vision comme un problème ancré

dans un espace à trois dimensions (où ? quoi ? comment ?) [Rao et Ballard, 1995; Garbay, 2000].

3.2.1 Les contraintes

Le problème de la vision soulève de multiples difficultés :

- on ne connaît pas *a priori* la chaîne de traitements qui permet de résoudre le problème de vision posé. Son choix dépend du contexte et des situations rencontrées. Le but n'est jamais défini de manière globale, bien au contraire des buts locaux surgissent selon les traitements effectués. D'autre part, la variabilité des apparences pour un même objet ou plusieurs objets dans une image suggère qu'il est impossible d'appliquer un traitement global à toute l'image. Aussi, le système doit en fonction du contexte adapter sa stratégie, les opérateurs ainsi que leurs paramètres suivant les informations locales récoltées.

- la spécification des entités, des objets, des classes, des sous-classes doit autoriser une grande variabilité intra et inter-classe. Ceci impose bien évidemment la gestion de l'incertitude au sein des systèmes.

- l'information d'intérêt est parfois concentrée sur quelques pixels, par exemple en détection, et elle est noyée dans une multitude de données. Aussi, il y a une nécessité de travailler de façon incrémentale, sûre et progressive de manière à accumuler des connaissances. L'objectif est de s'appuyer sur des résultats déjà obtenus pour poursuivre le traitement.

Dans les systèmes de vision, l'intégration de connaissances hétérogènes en grande quantité et la prise en compte de toutes ces contraintes, nous mènent directement vers des architectures distribuant la connaissance et permettant d'effectuer des traitements en parallèle et en concurrence [Sandakly, 1995]. Dans cette optique, une approche de type multi-agent constitue une réponse particulièrement adaptée [Ferber, 1995]. En effet, elle autorise une grande modularité et une adaptativité importante du système.

3.2.2 L'approche proposée

Un système de vision est défini comme un système exploratoire répondant à trois questions : quelle est l'information d'intérêt ? où se situe-t-elle ? et enfin comment est-il possible de l'extraire ? Pour répondre à ces questions et à l'ensemble des contraintes préalablement posées, un système de vision doit en permanence s'appuyer sur les informations collectées pour construire sa propre connaissance du domaine. L'objectif est de proposer de nouvelles zones de focalisation, des buts et des stratégies différentes en fonction des informations disponibles. Pour respecter cette

philosophie d'ancrage dans un espace multidimensionnel, nous avons effectué un ensemble de choix que nous justifierons au cours de ce chapitre. Mais nous pouvons déjà présenter les grandes lignes du système.

Espace du problème et agents situés

Nous proposons de centrer la conception de notre système sur la notion d'agents situés. Un tel agent est ancré dans un espace à trois dimensions : l'espace image (le "où"), l'espace des buts (le "quoi") et l'espace des méthodes (le "comment"). Une représentation du système est donnée sur la figure 3.1.

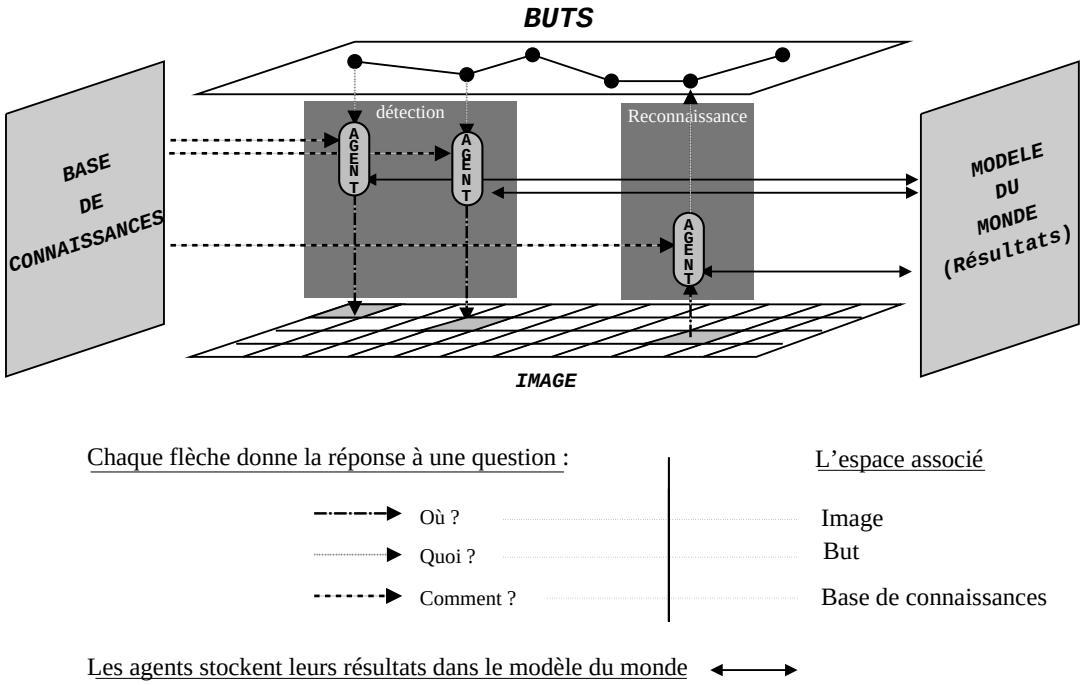


FIG. 3.1 – Vue générale du système multi-agent. Chaque agent est ancré dans un espace à 3 dimensions, constitué de l'espace image, l'espace des buts et l'espace des méthodes.

La base de connaissance autorise l'articulation entre différents espaces de représentation :

- but ↔ comment : la base de connaissances associe à chaque concept des méthodes qui seront ensuite ajustées.
- but ↔ où : ce lien est réalisé avec la spécification de stratégies de focalisation qui permettent pour un nouvel objectif de définir de nouvelles zones d'intérêt.

Chaque agent a une position donnée dans l'environnement, avec un but précis. Il travaille localement en utilisant les données de la base de connaissance pour produire des résultats partiels qui seront stockés dans le modèle du monde. Ce modèle

du monde est ensuite partagé par tous les agents.

Ce basculement entre les différents espaces de représentation est notamment vérifié pour les tâches de détection et de reconnaissance. Il est représenté sur la figure 3.1. En phase de détection un agent avec un but spécifique, sélectionne une méthode pour extraire une région d'intérêt. En phase de reconnaissance, un agent focalisé sur une région d'intérêt sélectionne une méthode pour réaliser un but.

Focalisation et Adaptation

Suivant la philosophie préconisée par la perception active et notamment par R. Bajcsy [Bajcsy, 1988], nous avons choisi une stratégie procédant par focalisations successives, c'est-à-dire un contrôle guidé par la recherche d'informations [Touboul, 1998]. Ainsi, lorsqu'un agent extrait une information d'intérêt, il peut demander la création d'un autre agent pour obtenir un complément d'information dans une région précise de l'image. En conséquence, le système élabore une stratégie en fonction de ce qu'il trouve à un instant donné, et on peut en ce sens le qualifier d'opportuniste. Selon le but recherché et le contexte, une méthode est sélectionnée dans la base de connaissances et ses paramètres sont ajustés localement par l'agent.

Incertitude et réseaux bayésiens

Enfin une caractéristique importante du système est l'incrémentalité. Ayant toujours à l'esprit que notre système se construit et vient enrichir sa connaissance au cours du temps, il est important de laisser des degrés de liberté dans son évolution sans figer les résultats. Nous avons donc mis en œuvre une stratégie incrémentale pour la mise à jour des confiances. Le système de vision fournit en permanence des résultats imparfaits et imprécis. Cette notion d'incertitude et de connaissance partielle sont des grandeurs importantes du système. Nous avons choisi une approche basée sur les réseaux bayésiens pour répondre à ce problème.

En conclusion

Nous avons présenté de façon très succincte notre approche. Celle-ci est basée sur une architecture multi-agent qui répond parfaitement au problème de vision posé comme un problème d'ancrage dans un espace à trois dimensions. Nous avons parlé de plusieurs grands principes avec la notion :

- d'agents situés ;
- de focalisation ;
- d'adaptativité ;
- d'incrémentalité ;
- et d'incertitude.

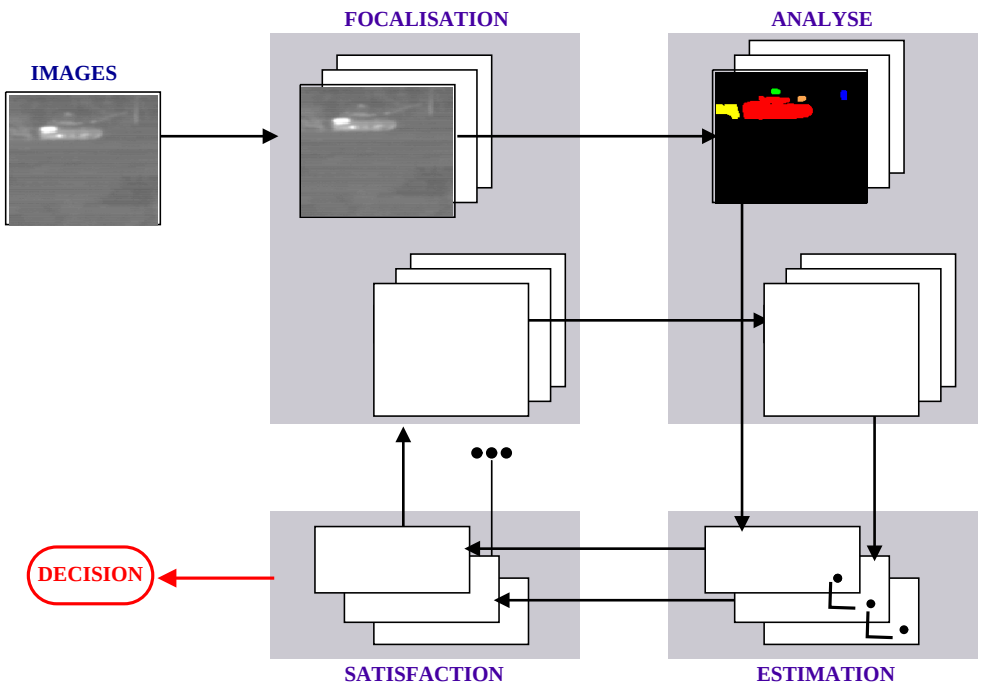


FIG. 3.2 – Approche proposée pour l'analyse de scènes. L'image ou les images sont les données d'entrée. Un comportement de focalisation permet de chercher les régions d'intérêt qui seront ensuite les données d'entrée des agents d'analyse tels que les agents détection. Pour chaque résultat, une confiance est calculée par le comportement d'estimation et celle-ci est propagée aux autres noeuds du réseau. Enfin, si un agent atteint une satisfaction (confiance) suffisante, il s'interrompt, sinon il relance son comportement de focalisation.

Ces choix se justifient et ont été inspirés par les qualités de systèmes existants. Au cours de l'exposé qui va suivre, nous proposons de recadrer plus précisément l'ensemble de ces choix dans un contexte bibliographique.

3.3 Focalisation et adaptation

Les mécanismes de focalisation et d'adaptation permettent d'améliorer les performances d'un système de vision. Ils permettent un comportement actif du système, en cherchant en permanence un compromis entre l'objectif poursuivi et l'information présente dans l'image [Bajcsy, 1988]. D'un point de vue méthodologique, la vision active tente ainsi d'améliorer les performances par rapport aux approches passives classiques. Elle cherche à élaborer des stratégies intelligentes en contrôlant par exemple, le choix des opérateurs, les valeurs de leurs paramètres, les zones d'observation, ou encore d'un point de vue matériel, la position des capteurs. La vision active est une approche qui apporte une solution aux problèmes de conception des systèmes de vision tels que leur sensibilité au bruit, leur faible réactivité et leur faible précision. C'est également une solution pour gérer la complexité des informations

en utilisant uniquement l'information pertinente et significative.

Cette philosophie a été particulièrement exploitée dans les systèmes dotés d'un contrôle décentralisé, comme par exemple le système SCHEMA proposé par Draper [Draper *et al.*, 1989]. Ce système fournit des interprétations partielles réalisées dans des régions de l'image qui peuvent à leur tour contraindre et guider le traitement ailleurs dans l'image. La conception de ce système repose sur la notion de schémas. Un schéma est un sous-système représentatif d'un concept comme un objet ou une partie d'objet. Il contient les connaissances descriptives et les stratégies de contrôle nécessaire à la reconnaissance de l'objet en question. Un schéma réalise ainsi le lien entre les espaces relatifs au "quoi" et au "comment". Chaque schéma dispose de son propre tableau noir ("blackboard") pour stocker les hypothèses et les informations collectées localement. Chaque schéma instancié active les différentes stratégies spécifiées (notion de concurrence) pour détecter les objets en localisant les régions de l'image, ou en mettant en œuvre les schémas associés aux sous-parties qui peuvent eux-mêmes activer d'autres schémas. L'ensemble de ces schémas activés communique via un tableau noir global qui selon son état courant permet de poursuivre l'analyse.

Ces deux notions que sont la focalisation et l'adaptation sont très liées. Il n'est possible d'adapter les paramètres et les stratégies d'un système que si au préalable une focalisation a été effectuée.

La focalisation et l'adaptation sont des principes qui s'intègrent complètement dans la notion d'agents situés. On retrouve cette approche dans le système proposé par A. Boucher [Boucher, 1999]. Ce système multi-agent est construit autour d'un modèle d'agent générique constitué de quatre comportements de base : perception (croissance de régions, suivi de contour), interaction (fusion de primitives), différenciation (interprétation de primitives), et reproduction (stratégie de focalisation des agents). Ce comportement de reproduction s'appuie dynamiquement sur les informations collectées dans l'image pour construire la stratégie d'exploration dans l'image ou dans une autre image de la séquence, suivant une philosophie opportuniste.

Nous venons d'illustrer les principes de focalisation et d'adaptation et leurs ancrage possibles dans les espaces "ou" et "comment". On peut continuer cette analyse en détaillant désormais chaque principe dans l'espace global de représentation. Chacun se décline à des niveaux différents dans un système de vision. On retiendra pour la focalisation les principes suivants :

- la focalisation spatiale : contrôle dans l'espace image
- la focalisation cognitive : contrôle dans l'espace des buts
- la focalisation opératoire : contrôle dans l'espace des tâches

Cette catégorisation fait ressortir une fois de plus, cet ancrage dans l'espace à trois dimensions qui s'avère très important. Aussi, nous proposons un système qui

puisse naviguer dans ces espaces librement. Chaque agent dispose comme donnée d'entrée d'un but et d'une zone de l'image. Ainsi, ancré dans une zone de l'image avec un but précis, il peut élaborer sa stratégie en fonction des informations disponibles. Et, sur ce point, on retrouve la notion d'adaptativité qui elle-même se décompose naturellement dans cet espace de représentation.

- l'adaptation image : le système évolue par focalisations successives et s'organise selon la nature des informations trouvées dans l'image ;
- l'adaptation cognitive : les buts sont construits de manière dynamique selon les informations collectées ;
- l'adaptation opératoire : les méthodes sont adaptées selon le but et le lien.

Cette description des principes de focalisation et d'adaptation permet de montrer le lien qui les unit. Il y a en effet une interaction importante entre ces deux notions. La focalisation apparaît en quelque sorte comme une étape indispensable à l'adaptation et ce quelque soit l'espace considéré. Par exemple dans l'espace image, la focalisation définit une fenêtre d'intérêt, et l'adaptation exploite les informations présentes dans cette fenêtre pour ajuster des paramètres.

La focalisation et l'adaptation au sein du système se font via une interaction permanente entre le système et son environnement. Suivant les informations disponibles, le système se construit et s'organise de façon à enrichir sa connaissance du domaine d'application.

3.4 Distribution et coopération

La distribution et la coopération sont deux principes que l'on peut associer naturellement. En effet, la distribution dans un ensemble de sous-espaces sous-entend la récolte de résultats partiels. Or l'objectif de la coopération est justement de coordonner les activités et de combiner les résultats de façon à réaliser la tâche de vision. Un état de l'art très intéressant a été proposé par C. Garbay [Garbay, 2000] et établi d'après les travaux de J. Hoc [Hoc, 1996]. Il distingue trois formes de coopération :

- la coopération confrontative est organisée selon une philosophie concurrente. Des informations de natures différentes sont collectées et fusionnées pour fournir une solution.
- la coopération augmentative procède suivant un principe de focalisation. C'est un assemblage des résultats partiels obtenus sur des ensembles disjoints de données, qui constitue la solution finale.
- la coopération intégrative décompose chaque problème de vision en sous-problèmes résolus de manière coordonnée. La solution finale est obtenue au terme de leur mise en œuvre.

Il est intéressant de faire le parallèle avec les trois espaces de représentation des agents pour souligner le lien entre la coopération et la distribution. L'analogie suivante apparaît :

coopération confrontative ↔ *comment* : on considère le problème dans sa distribution opératoire. Plusieurs méthodes peuvent être envisagées pour atteindre un but.

coopération augmentative ↔ *où* : on considère le problème comme distribué spatialement. Le système navigue dans l'espace image. Il travaille sur des zones de l'image puis fusionne les résultats.

coopération intégrative ↔ *quoi* : on considère le problème comme distribué fonctionnellement. Le système navigue dans l'espace des buts. Pour atteindre l'objectif final, il faut réaliser des tâches intermédiaires.

Cette notion de coopération inclut bien évidemment une dynamique collective et une distribution dans la conception. Dans les systèmes de vision en général, il y a toujours au moins un niveau de coopération [Gatepaille et Brunessaux, 1996].

3.4.1 Coopération confrontative

Ce type de coopération apparaît dans les systèmes de vision sous différentes formes suivant les informations manipulées. Il permet le traitement des données très hétérogènes qui peuvent être numériques (temps, images, signature radar...) ou symboliques par la mise en œuvre des techniques de la fusion de données. Le but de la fusion est de rassembler de manière cohérente les données disponibles de plusieurs sources pour l'examen d'un même problème afin d'améliorer la qualité des déductions.

La fusion numérique

Le problème de la fusion de données se pose dans la majorité des cas lors de l'analyse de scène réalisée à partir de sources multi-capteurs, ou d'images multi-bandes d'un même capteur. V. Lefevre pour un problème lié à la photo-interprétation, utilise des informations multi-capteurs (visible et infrarouge) extraites de différentes images pour réaliser une tache de vision [Lefevre, 1994].

On peut également fusionner plusieurs images d'un même capteur, mais à des instants différents. Avec fusion temporelle, il s'agit de raisonner sur des informations évoluant dans le temps (vidéo surveillance, détection d'alarmes, véhicules intelligents [Rombaut *et al.*, 2000]).

La fusion symbolique

Cette situation se rencontre classiquement lorsque l'on dispose au moins d'une source d'informations extérieures à l'image, comme par exemple, une carte, un atlas, etc. Il s'agit dans ce cas d'élaborer des schémas de prise de décision tirant profit de

la présence simultanée de ces deux types d'informations.
H. Moissinac-Massenat propose notamment une utilisation conjointe d'informations symboliques et de mesures numériques pour la prise de décision en photo-interprétation [Moissinac-Massenat, 1996].

Dans la fusion symbolique, on intègre également la notion de fusion de décision. L'objectif est d'utiliser de manière concurrente différents classifieurs fournissant des points de vue différents. Bloch propose à ce sujet une fusion horizontale. Un vecteur dont les coordonnées sont les degrés d'appartenance aux différents ensembles flous proposés, est associé à chaque point [Block et Maitre, 1994].

Le principe de la fusion confrontative est donc une manière de rassembler de façon cohérente les données disponibles pour améliorer la pertinence des hypothèses. Mais les philosophies de conception peuvent être différentes. On peut utiliser par exemple, des critères de pondération comme le propose A. Boucher [Boucher, 1999], ou utiliser des techniques plus sophistiquées comme la mise en concurrence dans les architectures distribuées, ou la compétition de compétences [Lefevre, 1994].

3.4.2 Coopération augmentative

L'objectif est de répartir la réalisation d'une tâche sur plusieurs entités. Cette approche a été mise en valeur dans le système multi-agent proposé par A. Boucher [Boucher, 1999]. En effet, plusieurs agents sont lancés dans l'image pour segmenter des objets de l'image. Lorsque deux agents se rencontrent et découvrent qu'ils travaillent sur la même composante, ils fusionnent les informations collectées, à savoir les structures d'informations, les statistiques, les primitives (contour et région), etc. Cette fusion ne concerne que les agents possédant la même spécialisation.

3.4.3 Coopération intégrative

Le principe de la coopération intégrative est de décomposer une tâche en plusieurs sous-tâches. Cette approche est utilisée notamment dans le système de vision MESSIE [Sandakly, 1995]. Pour atteindre un but de vision, les objets présents sont classés par ordre de facilité de détection. Cette notion de facilité est une notion comparative qui se base sur la description géométrique des objets et leur apparence radiométrique. Par exemple, un objet sera plus facile à détecter si ses dimensions sont grandes. La stratégie utilisée dans MESSIE est de se focaliser d'abord sur les objets les plus évidents de l'image (les grandes structures) pour ensuite affiner l'analyse et en identifier les détails.

Nous retiendrons pour notre système, ce principe de construction d'un chemin pour atteindre un but qui s'appuie sur la notion de facilité de détection. Il est particulièrement adapté à notre problème, puisque en phase de détection, l'information d'intérêt est parfois concentrée sur quelques pixels. Aussi, il paraît judicieux de

garder cette philosophie pour rechercher les éléments contextuels. Pour trouver les véhicules présents dans une image, il est sans doute judicieux de chercher les routes et d'exploiter les relations de proximité ensuite. C'est une manière d'accumuler des informations complémentaires.

3.4.4 Conclusion

Nous avons présenté différentes formes de coopération au sein des systèmes de vision. Elles se déclinent dans des espaces différents (images, buts et méthodes), mais toujours dans le but d'accumuler des informations supplémentaires. Chaque méthode de coopération présente un intérêt qui se situe à différents niveaux de conception du système de vision. En conséquence nous avons retenu chacune de ces possibilités pour construire notre système, notamment :

- pour localiser un objet nous utiliserons des informations de natures différentes pouvant éventuellement provenir de capteurs différents.
- on retrouvera la notion de coopération augmentative proposée par A. Boucher.
- enfin, comme nous l'avons déjà précisé nous chercherons toujours, en détection, à extraire les éléments significatifs de l'image avant d'affiner l'analyse. Selon le principe de la focalisation intégrative, il y a un ordre dans les buts à atteindre qui conditionne ensuite le choix des zones d'intérêt (où) et des méthodes de traitement (comment).

Le point fondamental caractérisant l'ensemble de ces coopérations qu'elles soient confrontative augmentative, ou intégrative est l'accumulation d'informations dans le but de prendre une décision.

3.5 L'incertitude

Au cours du processus de vision, nous sommes contraints de prendre des décisions. Ces choix s'inscrivent aussi bien dans un cadre passif (recherche arborescente), que dans un cadre actif (choix d'une action pour acquérir de nouvelles données). Pour effectuer ces choix, nous devons à tout moment estimer notre confiance dans les hypothèses courantes. Une hypothèse concerne le type de la cible, sa position et son attitude éventuellement. La confiance dépend de très nombreux paramètres tels que :

- la qualité du traitement ;
- le choix des opérateurs ;
- la présence de descripteurs significatifs pour les objets ;
- la connaissance sur l'environnement.

Elle intègre des données de natures très différentes telles que les informations statistiques, géométriques ou cinématiques par exemple. Notre tâche consiste à for-

maliser un raisonnement nous permettant de localiser puis de reconnaître une cible dans un environnement inconnu. Pour cela différentes méthodes existent telles que les approches probabilistes [Rimey, 1993], possibilistes [Nifle, 1998], la théorie de Dempster-Shafer [Lefèvre *et al.*, 1996; Nigro *et al.*, 2000], ou encore les approches floues [Mees, 2000]. Ces formalismes modélisent l'incertain, en associant au calcul numérique un raisonnement. Ils permettent d'aborder le problème de la dynamique des croyances par l'intermédiaire de mécanisme d'inférence de façon à tirer des conclusions provisoires [Fabiani, 1996].

Comme le précisent très justement F. Bremond et M. Thonnat, *les approches les plus communément utilisées sont probabilistes... Dans le cadre de l'interprétation de scènes... elles offrent un cadre strict et rigoureux applicable à tout type d'incertitude* [Brémond et Thonnat, 1996].

Les graphes constituent un moyen graphique et efficace pour représenter les relations qui existent entre les différents états possibles du système : les liens de causalité, de dépendance, les relations temporelles, etc. Ils sont également un outil adapté pour la structuration des connaissances autorisant notamment la propagation des informations. Aussi, nous avons porté notre choix sur ce type de représentation.

Nous avons choisi de présenter les deux formalismes graphiques et probabilistes les plus utilisés et leur cadre d'application. Nous commencerons par décrire complètement l'approche par réseaux bayésiens que nous avons choisie. Nous montrerons en quoi elle est bien adaptée à notre problème de vision. Puis nous présenterons plus succinctement le formalisme mathématique des modèles de markov cachés, sachant qu'il présente des points communs avec les réseaux bayésiens.

3.5.1 Les réseaux bayésiens

Un réseau bayésien [Pearl, 1988] définit les relations causales entre des variables aléatoires qui peuvent avoir un nombre discret d'états possibles ou qui peuvent être continues selon une loi de Gauss par exemple.

Le principe

Un réseau bayésien est défini par :

- un graphe acyclique orienté G avec G=(N,A) où N est l'ensemble des noeuds du graphe et A l'ensemble des arcs du graphe ;
- un espace probabilisé fini (Ω,Z,p). Ω définit un univers, Z une tribu de Ω et p une loi de probabilité ;
- un ensemble de variables aléatoires définies sur (Ω,Z,p), correspondant aux noeuds du graphe :

$$P(N_1, N_2..., N_n) = \prod_{i=1,...n} p(N_i/C(N_i))$$

sachant que $C(N_i)$ représente les causes de N_i dans le graphe, c'est à dire ses parents.

La première étape dans un tel réseau est de spécifier le graphe et les relations de dépendances entre les variables aléatoires. La figure 3.3 donne une illustration de la représentation graphique d'un réseau bayésien constitué de six noeuds dont les états sont binaires (vrai, faux). Dans cet exemple, la distribution du fils C conditionné par ses parents A et B est représentée par une table de probabilités conditionnelles.

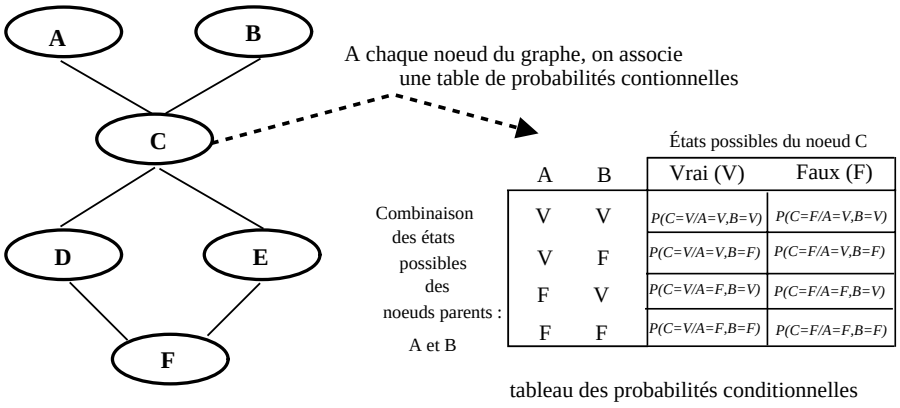


FIG. 3.3 – Exemple de réseau bayésien

L'étape suivante consiste à fixer la valeur des variables du réseau en question. Par exemple pour une variable booléenne, les états possibles sont Vrai ou Faux. La dernière étape consiste à propager l'influence des variables entre elles. Il existe deux types de probabilités calculables dans un réseau bayésien à savoir les probabilités *a priori* et les probabilités *a posteriori*. Mais, quelle que soit la probabilité recherchée, le mécanisme d'inférence utilisé pour la propagation des informations est défini par la loi de Bayes. Un exemple est présenté en annexe A.

Soient H et X deux variables aléatoires et H l'unique fils de X :

$$P(H/X) = \frac{P(H,X)}{P(X)} = \frac{P(X/H).P(H)}{P(X)}$$

avec $P(X) = P(X/H).P(H) + P(X/-H).P(-H)$

$P(H/X)$ représente la probabilité *a posteriori* d'un état possible de l'hypothèse H, sachant l'état de l'observation X,

$P(X/H)$ représente la vraisemblance d'avoir un état pour l'observation X connaissant l'état de l'hypothèse,

$P(H)$ et $P(X)$ représentent respectivement les probabilités *a priori* d'avoir une hypothèse donnée et une certaine observation.

Le processus de décision s'effectue par le calcul du maximum de vraisemblance. Il existe des algorithmes de propagation exacts ou approximés. On pourra à ce sujet se référer aux travaux de Spiegelhalter et Lauritzen [Spiegelhalter et Lauritzen, 1990], de Jensen *et al.* [Jensen *et al.*, 1990], ou encore au livre proposé par Becker et Naïm [Becker et Naïm, 1999].

Caractéristiques de l'approche

Les réseaux bayésiens sont fondés sur la théorie des probabilités et la loi de Bayes qui permet de calculer les probabilités *a posteriori*, connaissant les probabilités *a priori*. Cette approche nécessite la spécification d'un modèle probabiliste complet. La représentation des dépendances entre les variables aléatoires est particulièrement bien adaptée dans le cas d'observations contenant des incertitudes que l'on sait modéliser.

Le mécanisme d'inférence offre une grande flexibilité. Il permet d'instancier indifféremment les variables associées aux observations et celles liées à leurs causes. De plus la propagation de l'information est bidirectionnelle. Elle autorise un raisonnement dans le sens de la causalité, mais également dans le sens inverse.

Les réseaux bayésiens répondent bien à la difficulté de notre problème. Ils permettent

- de représenter graphiquement notre connaissance *a priori* sur l'environnement sous forme d'un réseau. Elle se situe dans la structure du graphe, mais également dans les tables de probabilités associées aux noeuds.
- de prendre en compte de nouvelles observations et de propager leur influence à l'ensemble du réseau.
- de quantifier les incertitudes dans le cadre rigoureux de la théorie des probabilités.

Applications

Les réseaux bayésiens ont été appliqués à de nombreux travaux liés à l'interprétation de scènes. Ainsi Rimey propose un système sélectif de vision (TEA-1) dont l'objectif est de déterminer les zones de la scène les plus intéressantes à explorer [Rimey, 1993]. TEA-1 est un système qui est basé sur la collecte séquentielle d'informations relatives à des tâches spécifiques. Ces tâches concernent l'interprétation d'une scène (sa composition) ou la détection d'un objet. Le système n'exécute que les actions nécessaires pour répondre au problème posé. Ce système utilise les réseaux bayésiens pour représenter la connaissance. Deux réseaux ("Part-of" et "Is-A") donnent une description de la scène sous forme de relations de composition et de spécialisation. Ils décrivent les relations entre les sous-ensembles de la scène. Un troisième réseau ("expected area net") décrit les relations géométriques entre

les objets et leurs localisations spatiales relatives dans l'espace considéré. Et enfin un quatrième réseau ("Task net") regroupe les informations utiles, en provenance des précédents réseaux, pour effectuer chaque tâche. La meilleure action est choisie à chaque instant de façon à optimiser le compromis coût/utilité. Chaque situation peut potentiellement impliquer plusieurs décisions intermédiaires, incluant des choix et donc des interactions avec l'environnement. TEA-1 cherche le "meilleur" enchaînement d'actions possible pour résoudre un problème donné. La stratégie mise en œuvre dépend de plusieurs paramètres et notamment des probabilités courantes, de l'information utilité, et de l'information coût. Ce système se situe à l'interface entre la représentation de la connaissance et la prise de décision.

Cette représentation est largement utilisée. Toujours dans le domaine de l'interprétation de scène et la reconnaissance d'objets, il est possible de citer les travaux de D. Djian [Djian, 1997], de S. Jonquières [Jonquières, 2000] et de Marchand et Chaumette [Marchand et Chaumette, 1998]. Dans le domaine de la surveillance du trafic routier, on peut citer les travaux de Buxton et Gong [Buxton et Gong, 1995] ainsi que Huang et Russel [Huang et Russell, 1998]. A. Stassopoulou et T. Calli propose l'utilisation de réseau bayésien pour résoudre la problème de la détection de bâtiments [Stassopoulou et Caelli, 2000].

3.5.2 Les chaines de Markov cachées

Les modèles de Markov cachés (*Hidden Markov Model* ou HMM) sont une extension des chaînes de Markov. Une chaîne de Markov décrit des processus de Markov discrets par un automate à état finis (Figure 3.4).

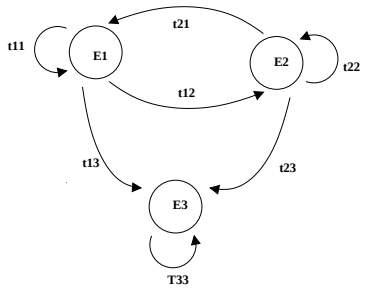


TABLE DES TRANSITIONS			
Xt+1 / Xt	E1	E2	E3
E1	t11	t12	t13
E2	t21	t22	t23
E3			t33

Fig. 3.4 – Exemples de chaîne de Markov

Un modèle de Markov est un modèle graphique : les noeuds représentent les variables aléatoires et les arcs représentent les transitions entre états. Dans une chaîne de Markov, les transitions sont probabilistes et elles sont représentées sous forme de matrice. Chaque élément t_{ij} de la matrice spécifie la probabilité de passer de l'état E_i à l'état E_j. Pour chaque ligne, la somme des éléments a_{i*} est égale à 1. Graphiquement ceci se traduit par la somme des arcs émergents pour un noeud donné est égale à 1. Pour chaque état, les évènements de transition sont complémentaires.

L'inférence

Les processus discrets vérifient l'hypothèse de Markov à l'ordre k. Soient f le processus considéré, X la variable d'état (à valeur dans E1, E2, E3 pour la figure 3.4), on a :

$$f_{t+1/t,...,0}(X_{t+1}/X_t, ..., X_0) = f_{t+1/t,...,t-k+1}(X_{t+1}/X_t, ..., X_{t-k+1})$$

à l'odre 1, cette formulation mathématique s'écrit :

$$f_{t+1/t,...,0}(X_{t+1}/X_t, ..., X_0) = f_{t+1/t}(X_{t+1}/X_t)$$

Ainsi, lorsque le présent est connu, le passé n'a plus d'influence sur le futur.

Un modèle de Markov caché est une chaîne de Markov du premier ordre avec des états non directement observables. En parallèle de la table de transition, il faut donc spécifier la table des observations (figure 3.5)

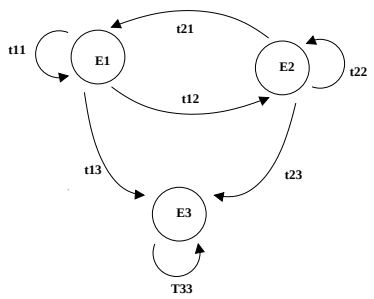


TABLE DES TRANSITIONS			
Xt \ Xt+1	E1	E2	E3
E1	t11	t12	t13
E2	t21	t22	t23
E3			t33

TABLE DES OBSERVATIONS			
	E1	E2	E3
O1	P1	P2	P3
O2	1-P1	1-P2	1-P3

FIG. 3.5 – Exemples d'un modèle de Markov caché

Pour chaque état, il y a deux observations possibles O1 et O2 avec des probabilités associées fonction de l'état. Pour décrire complètement un modèle, il faut décrire les états, les observations pouvant être obtenues, la matrice des observations, la matrice des transitions et la distribution initiale des états.

Caractéristiques de l'approche

La propagation de l'information dans le système se fait selon l'hypothèse markovienne. Ce formalisme est intéressant dans le cadre de la reconnaissance d'objet. Le fait que l'état réel ne soit pas connu exactement reflète bien la réalité d'un environnement observé à travers des capteurs. Cependant le modèle des incertitudes associées aux capteurs par exemple, est limité à une distribution de probabilités entre état et observation. Il paraît mieux adapté aux applications où la notion de temps est plus explicite.

La représentation sous forme d'un réseau bayésien est directe si on fait l'hypothèse que les noeuds cachés sont discrets. Pour cela, il faut associer à chaque états possibles, une observation. Un exemple est présenté sur la figure 3.6 qui distingue deux états possibles E1 et E2 et deux observations associées O1 et O2.

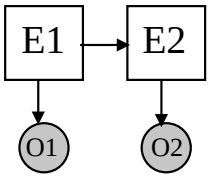


FIG. 3.6 – Représentation d'un modèle de Markov caché sous forme d'un réseau bayésien

Les applications

La représentation de l'incertitude sous forme d'un modèle markovien a été utilisée dans plusieurs domaines, notamment dans le cadre de l'apprentissage et spécifiquement de l'apprentissage par renforcement (Q-learning) [Kaelbling *et al.*, 1996; Keerthi et Ravindran, 1994; Dearden *et al.*, 1998]. H. Moissinac l'a appliqué au problème de l'interprétation de scènes urbaines [Moissinac *et al.*, 1995]. W. Burger and B. Bhanu propose une approche pour la détection d'objets [Burger et Bhanu, 94].

3.5.3 En conclusion

Les graphes permettent de représenter les relations entre les états possibles. Les méthodes de calcul associées sont traitées dans la théorie des graphes. Ils constituent, entre autres, un outil adapté à la représentation des relations de causalité, des relations temporelles. Il est possible d'associer aux arcs d'un graphe des conditions de transition, des notions de durée, de probabilité de transition, ou de degré de croyance pour représenter l'incertitude d'une relation de causalité.

Le formalisme bayésien nécessite la définition complète du modèle probabiliste mais il offre un cadre rigoureux [Fabiani, 1996]. Notre choix s'est donc porté sur ce formalisme. Nous avons construit les réseaux avec l'outil "Bayes Net Toolbox" (cf. annexe A). Ces réseaux sont utilisés pour propager les incertitudes du système dans le but de renforcer ou diminuer des hypothèses.

3.6 Incrémentalité et apprentissage

Nous avons déjà abordé, au cours de cet exposé, le problème de l'incrémentalité dans les systèmes de vision. Nous pouvons reparler du système MESSIE proposé par Sandakly [Sandakly, 1995]. Ce système exploite, dans un contexte multi-capteur, les

informations “faciles” à identifier telles que les grandes structures de l’image, pour ensuite affiner l’analyse. Le système se construit progressivement en fonction des informations disponibles. Il y a une incrémentalité dans ce processus de collecte d’informations.

Bien évidemment cette incrémentalité se retrouve dans quasiment tous les systèmes de vision car elle est intimement liée au contrôle lorsque l’on parle d’incrémentalité sur les actions. Dans les réseaux bayésiens notamment, cette incrémentalité existe dans la mise à jour des confiances. Rimey propose d’enchaîner des actions suivant un processus de coût/utilité [Rimey, 1993]. Il se pose la question de savoir quelle est la meilleure action à effectuer à un instant donné. Ce principe nous sera particulièrement utile en phase de reconnaissance. En effet, certaines informations redondantes entachées de bruit, dissimulent des informations pertinentes, alors qu’une autre information sera très discriminante. L’objectif sera d’utiliser un critère de coût pour connaître l’attribut le plus significatif.

D’autres formes un peu plus atypiques existent. On pourra prendre l’exemple du système CITE proposé par Dillon et Caelli qui intègre les notions d’incrémentalité et d’apprentissage dans une même boucle [Dillion et Caelli, 1998]. Ce système est doté d’une base de connaissances qui fournit une description de la scène. La mise à jour de cette connaissance est effectuée suivant un processus incrémental supervisé. Cette approche permet de construire et d’améliorer la connaissance au cours du temps. Il y a alors un lien direct entre l’incrémentalité et l’apprentissage “on-line”. Ce mode d’apprentissentage vient complèter l’approche traditionnelle qui consiste à donner une modélisation *a priori* des objets. L’intérêt de l’apprentissage “on line” réside dans son caractère dynamique et incrémentale. Il peut ainsi autoriser une mise à jour des connaissances au sein du système.

Cette forme d’apprentissage dynamique est très intéressante car elle permet d’établir un lien entre l’adaptation et l’incrémentalité qui sont des principes importants de notre approche. La modélisation utilisée peut ainsi progressivement se détacher d’un *a priori* souvent partiel et distant de la réalité.

3.7 Conclusion

Au cours de ce chapitre nous avons voulu faire ressortir des points fondamentaux des systèmes de vision. Ils sont représentés par un ensemble de mots clés :

- focalisation et adaptation
- distribution et coopération
- incertitude
- incrémentalité et apprentissage

Pour chaque terme nous avons illustré leur utilisation dans un certain nombre de systèmes de vision. L'objectif était bien évidemment de justifier à chaque étape nos choix de conception et de présenter les systèmes qui nous inspiré.

Nous proposons une architecture multi-agent (distribution et coopération) basée sur la mise en œuvre d'agents situés dans l'espace image, dans l'espace de buts et dans l'espace des méthodes. Cet ancrage répond aux principes de focalisation et d'adaptation. Le système propose une conception coopérative de la résolution de problème. Il établit sa stratégie de résolution en fonction des informations disponibles et pertinentes (incertitude) et des règles de focalisation (incrémentalité). La gestion des incertitudes se fera sous forme de réseaux bayésiens.

Le mot apprentissage qui apparaît n'est pas encore véritablement inclus dans le système, si ce n'est au travers d'une mise à jour de la base de connaissance. Mais, il reste important et fera partir des futurs objectifs pour notamment fixer les paramètres des réseaux bayésiens.

Suite à cette analyse conceptuelle de l'approche nous allons désormais en donner une description détaillée.

Chapitre 4

Description du système

Sommaire

4.1 Introduction 73
4.2 Principe de fonctionnement 73
4.3 Plate forme multi-agent 75
4.4 Description fonctionnelle du système 77
4.5 Représentation des informations 83
4.6 Comportements des agents 92
4.7 Conclusion 106

Résumé

Dans ce chapitre, nous proposons une architecture multi-agent pour résoudre le problème de la détection de la reconnaissance et de l'identification de cibles (DRI). Après une description de la plate-forme utilisée, nous présenterons la philosophie générale de résolution du problème en détaillant les cycles de contrôle intra et inter-image. La stratégie de résolution mise en œuvre par une population d'agents, qui explore de façon incrémentale l'espace image, avec des objectifs différents.

Chaque agent utilise les éléments spécifiés dans la base de connaissances pour réaliser son objectif. Les résultats extraits et validés sont quant à eux stockés dans le modèle du monde. Nous présenterons successivement ces deux structures de données.

Nous terminerons ce chapitre en détaillant les comportements mis en œuvre par les différents agents du système.

4.1 Introduction

Nous proposons une architecture décentralisée pour répondre au problème de la détection de la reconnaissance et de l'identification de cibles (DRI). Cette architecture s'articule autour de trois composantes principales :

- les agents constituent la partie "active du système". Ils explorent l'image incrémentalement. Ils sont ancrés dans l'espace de l'image, dans l'espace des buts et dans l'espace des méthodes.
- la base de connaissances fournit les connaissances descriptives et opératoires nécessaires aux agents. Elle précise l'espace image considéré et donne une description des buts et des méthodes associées. C'est un modèle de scène suivant la terminologie employée au chapitre 2.
- le modèle du monde regroupe l'ensemble des informations extraites par les agents. Il donne une représentation de la scène qui s'enrichit au cours du temps. C'est un modèle d'observation.

En exploitant la base de connaissances, les agents vont explorer l'espace image pour détecter, reconnaître et identifier les objets présents. Chaque agent dispose d'une autonomie pour accéder aux données, et élaborer sa propre stratégie. La planification est prise en charge localement.

Dans une première partie, nous poserons le principe de fonctionnement de notre système. Ensuite, nous présenterons la plate-forme utilisée et les différents éléments qui la composent à savoir, le modèle d'agents générique et l'administrateur. Dans une deuxième partie nous présenterons notre système et son schéma de contrôle. Dans les parties suivantes, nous détaillerons davantage le système en présentant les différents agents, les informations qu'ils manipulent et les comportements qu'ils adoptent.

4.2 Principe de fonctionnement

Le système d'analyse d'images s'articule autour de la mise en œuvre d'une population d'agents situés dans la séquence d'images. Chaque agent exploite une région de l'image afin d'extraire une information d'intérêt relative à un concept.

Qu'est-ce qu'un agent ?
Un agent est un processus informatique autonome qui travaille dans une région de l'image avec un objectif qui lui est propre. Un objectif correspond à un élément de la liste des concepts.

Qu'est-ce qu'un concept ?
Un concept est un élément de la scène. Par exemple, dans l'image 4.1, on définit quatre concepts : les routes, les champs, la végétation et les véhicules. L'ensemble de

ces concepts est décrit d'un point de vue opératoire dans la base de connaissances. La population d'agents va utiliser cette connaissance de façon à extraire le support

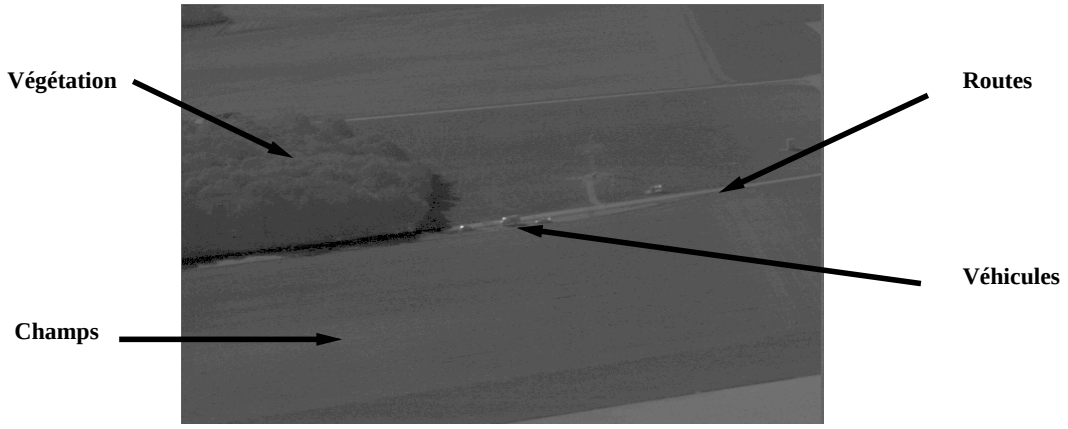


Fig. 4.1 – Les concepts de la scène : on distingue les routes, la végétation, les champs et les véhicules.

des objets (définition du mot support dans le paragraphe 4.4.1) et à les décrire, en exploitant progressivement la séquence d'images. Le but est de tirer profit des informations accumulées au cours du temps pour affiner l'analyse et les résultats.

Il est important de souligner que la recherche d'un support associé à un concept est effectuée par un groupe d'agents. Plusieurs agents par exemple participeront à l'extraction d'un support de type champ. Ainsi, il est possible d'augmenter la distribution spatiale et opératoire des traitements dans l'espace image. Les agents travaillent de façon concurrente et parallèle.

Une représentation très schématique du système est donnée sur la figure 4.2.

Plusieurs agents travaillent à un instant donné dans l'image, ou la séquence d'images, avec un objectif propre. Ils peuvent communiquer entre eux ou avec d'autres modules du système tels que l'administrateur, le modèle du monde ou encore la base de connaissances.

Chaque agent utilise les connaissances spécifiées dans la base de connaissances pour réaliser son objectif. Il stocke l'ensemble des résultats obtenus dans le modèle du monde. Enfin, il peut demander à l'administrateur la création d'un nouvel agent, ayant un rôle ou un objectif différent pour approfondir l'analyse.

La logique globale de résolution du problème s'appuie sur la mise en œuvre d'agents qui diffèrent tant par leur rôle, leur comportement que par les informations manipulées et produites. Le but est de distribuer les traitements utilisés.

Nous détaillerons dans la suite de ce chapitre l'ensemble de ces modules à sa-

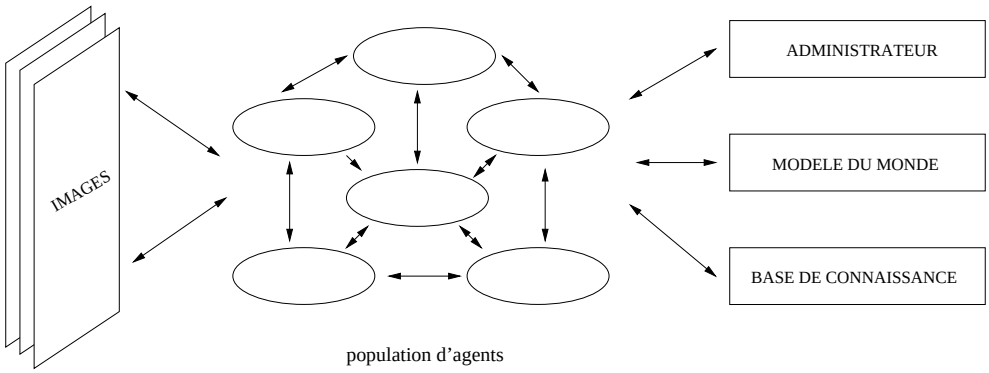


FIG. 4.2 – L'architecture du système multi-agent. À chaque instant, une population d'agents travaille dans la séquence d'images. Chaque agent utilise la base de connaissances pour sélectionner les méthodes et stocke les résultats fournis dans le modèle du monde. Un agent peut, de plus, demander à l'administrateur la création d'un nouvel agent.

voir l'administrateur, la base de connaissances, le modèle du monde ainsi que la population d'agents. Toutefois, afin de faciliter la compréhension du système nous décrirons, en premier lieu, la plate-forme utilisée et le schéma de contrôle au travers des différents cycles d'analyse relatifs à l'image puis à la séquence d'images.

4.3 Plate forme multi-agent

Une communauté d'agents ne peut exister sans un ensemble de modèles plus ou moins complexes définis au sein d'une plate-forme. Cette structure précise les éléments fondamentaux tels que l'architecture matérielle et logicielle ou encore les modes de communication permettant aux agents de communiquer entre eux et avec les données.

4.3.1 Présentation de la plate-forme

La plate-forme utilisée est reprise du système d'A. Bouillot auquel on a adjoint une gestion dynamique des agents [Bouillot, 1999]. Elle est composée de deux éléments : une structure d'agent dotée de moyens de communication et un superviseur indépendant appelé administrateur.

Les différentes fonctionnalités proposées par cette structure sont synthétisées figure 4.3 selon le plan proposé par le groupe de travail ASA (Architecture de Systèmes d'Agents) commun à l'AFIA[1] et au GDR I3[2].

[1]Association Française d'Intelligence Artificielle.
[2]Groupement De Recherche "Information Interaction Intelligence".

Nom de la plate-forme	Aucun.
Date et contexte de création	Développée en 1998 par A. Bouillot et C. Perez (DGA) comme outil pour différents travaux de recherche.
Degré d'avancement	Version 1.0
État actuel	Version stable.
Configuration matérielle	Fonctionne sur un ensemble de stations UNIX (dont une machine 32 processeurs).
Environnement logiciel	⋆ Programmation C++. ⋆ Bibliothèque de *threads*. ⋆ Interface TCL/TK.
Documentation	Rapport de stage [Bouillot, 1999]. Il n'existe pas de documentation spécifique disponible actuellement.
Application	⋆ Application à la robotique mobile [Dalgalarrondo, 2001]. ⋆ Mise en œuvre d'une chaîne de traitement d'image [Ropert, a paraitre]. ⋆ Application à la détection, la reconnaissance et l'identification de cibles.[Ealet *et al.*, 1999; Ealet *et al.*, 2000]
Architecture générale	Chaque agent est un processus indépendant (*thread*). Ces agents communiquent par envoi de messages suivant un protocole asynchrone. Tout agent peut demander la création d'un autre agent ou décider de quitter le système. Respectivement les messages d'activation ou de destruction sont envoyés à l'administrateur.
Modèle d'agent	Le modèle d'agent générique comprend toutes les fonctionnalités de communication mais aucune donnée particulière concernant les compétences de l'agent, la représentation de tâches extérieures ou la représentation des compétences des autres agents du système. Mis à part les fonctions de communication, l'agent est donc une structure vide.
Outils de mise au point	Pas d'outil de mise au point, en revanche il existe une interface permettant de conserver une trace du déroulement du système.
Test	Aucun test.
Disponibilité	Non disponible.

FIG. 4.3 – Description de la plate-forme multi-agent

4.3.2 Le modèle d'agent

Le modèle d'agent est défini par une classe générique encapsulant les mécanismes de communication afin d'autoriser les échanges d'informations au sein de la population. Une description de ce modèle est donnée en annexe B.1. Le protocole de communication choisi est basé sur l'échange des messages. Chaque agent archive ses messages dans une structure de type boîte aux lettres. Des requêtes ou des questions peuvent ainsi lui être adressées. En résumé le mécanisme de communication gère :

- le stockage des messages ;
- l'envoi des messages ;
- la lecture des messages ;
- l'attente de messages particulier.

Dans la conception du système, tout agent du système héritera de ce modèle générique incluant les mécanismes de communication.

4.3.3 L'administrateur

L'administrateur assure la coordination entre les différents acteurs du système. Les agents se déclarent auprès de l'administrateur qui les enregistre et leur attribue une fiche d'information.

Son rôle est de :
- recenser la population d'agents ;
- attribuer aux agents les ressources nécessaires à leur fonctionnement (allocation d'espace mémoire) ;
- déclencher l'activation d'un processus informatique de type thread pour chaque agent ;
- fournir un repère temporel commun.

D'un point de vue informatique, l'administrateur est déclaré comme une variable globale du système. Il est défini en C++ comme un objet qui est connu de tous les agents. L'administrateur est défini dans le programme par cinq structures différentes (cf. annexe B.2) :

- les fiches d'informations relatives aux agents enregistrés à un instant t (cf. annexe B.3) ;
- les variables liées à la population d'agents actifs ;
- l'historique des messages envoyés et reçus par les agents ;
- les variables liées à la gestion des threads ;
- les structures temporelles.

L'administrateur dispose d'une fiche d'informations spécifiant les caractéristiques techniques de chaque agent. Ainsi, un agent peut interroger cette base d'informations pour connaître son environnement.

4.4 Description fonctionnelle du système

Le but de ce paragraphe est de donner une description fonctionnelle de notre système. Après une spécification des notations utilisées, nous détaillerons les processus d'analyse intra-image et inter-image.

4.4.1 Les notations adoptées

Les agents du système manipulent des informations de natures différentes au cours de l'exploration. Nous distinguons principalement cinq structures issues de mesures (les trois premières structures) ou connues *a priori* (les deux dernières) :

- la zone d'intérêt Z est un élément de $\mathcal{Z}$ l'ensemble des zones possibles. Elle est définie par les deux couples de coordonnées associés à un rectangle englobant ;

- l'échantillon E est un élément de $\mathcal{E}$ l'ensemble des échantillons, il est représenté par une région dans l'image et caractérise une partie d'objet, une région élémentaire. Une étiquette, une croyance et un ensemble de paramètres descriptifs sont associés à chaque échantillon ;

- le support S est un élément de $\mathcal{S}$ l'ensemble des supports. Il est également représenté par une région dans l'image mais il caractérise l'objet dans sa globalité. Une étiquette, une croyance, des paramètres descriptifs et une liste d'échantillon décrivent chaque support ;

- l'objectif O est un élément de $\mathcal{C}$ l'ensemble des concepts. Il est initialement fixé par l'utilisateur. Il est mis à jour dynamiquement puis il est fourni par l'agent à ses fils ;

- le concept C est un élément de $\mathcal{C}$ l'ensemble des concepts. Il caractérise tous les objets d'un même type. Il est décrit par une étiquette et une liste de supports.

Une illustration des structures issues des mesures image, est présentée sur la figure 4.4 :

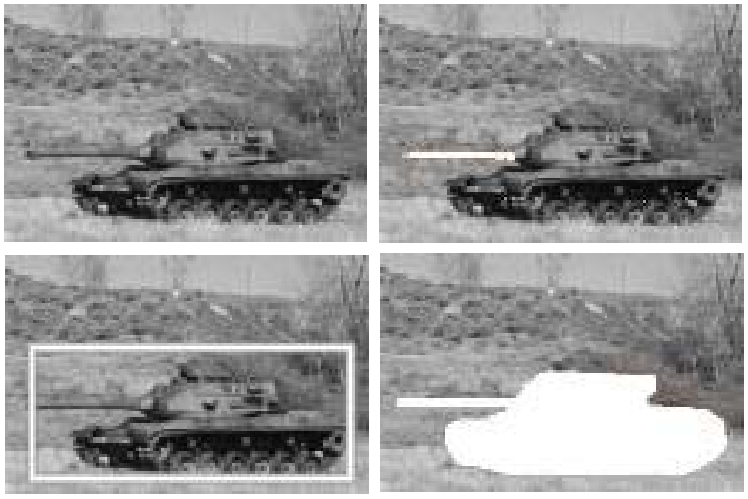


Fig. 4.4 – Les structures issues de l'image : en haut a gauche, on observe l'image de référence, en dessous le cadre blanc montre la zone d'intérêt, enfin à gauche se situent les illustrations de la recherche d'échantillons (en haut) et du support de l'objet (en bas). Ces structures représentées en blanc, correspondent à des régions de l'image caractérisant repectivement une région de l'objet ou un objet dans sa globalité.

4.4.2 Le cycle d'analyse intra-image

Le cycle d'analyse intra-image est composé de deux boucles distinctes :

- la première correspond à l'exploration spatiale de l'image et à l'exploration symbolique des concepts.
- la seconde permet d'exploiter des informations de natures différentes afin de valider ou de réfuter des hypothèses. Elle procède par fusion d'informations.

Nous présenterons successivement ces phases dans les deux paragraphes suivants.

L'exploration spatiale et symbolique

Cette exploration de l'image se décompose en trois parties :

- la première étape consiste à rechercher des échantillons E dans des zones de l'image c'est-à-dire des points d'ancrage possibles pour un concept considéré (route, champs, véhicule, ...). Concrètement on recherche les régions qui vérifient un ensemble des clauses spécifiées dans la base de connaissances (une description détaillée de cette structure est présentée dans le paragraphe 4.5.1). C'est la phase de localisation.

- la seconde étape exploite ces échantillons pour extraire les supports des concepts, par le biais notamment de techniques basées sur la croissance de régions. C'est l'étape de détection (attention ici le mot détection ne correspond pas au mot détection au sein de la fonction DRI).

- et enfin la dernière phase utilise les informations disponibles pour déterminer de nouvelles zones d'intérêt dans l'image et progresser dans l'espace des concepts. C'est l'étape de focalisation.

L'itération de ces trois phases se poursuit jusqu'à ce que toutes les règles de focalisation spécifiées dans la base de connaissances aient été explorées.

Nous avons fait le choix de distinguer les étapes de localisation et de détection pour deux raisons majeures. D'une part, il est souvent plus facile de trouver dans l'image un élément caractéristique d'un objet plutôt que le support complet. D'autre part, lorsque l'on a localisé les objets dans l'image, il est possible d'adapter localement les opérateurs et les paramètres afin d'améliorer les performances de segmentation suivant un principe d'incrémentalité.

La réalisation de chaque étape du processus est prise en charge localement par un type d'agent. Aussi, le cycle réalisant l'exploration spatiale de l'image, par focalisations successives, peut être schématisé par la figure 4.5. Chaque agent demande la création d'un nouvel agent à l'administrateur qui réalise alors la procédure d'enregistrement et qui active un nouveau processus informatique de type "thread".

L'agent Localisation cherche les échantillons (régions élémentaires) dans l'image. Pour chaque échantillon trouvé, il demande à l'administrateur la création d'un agent

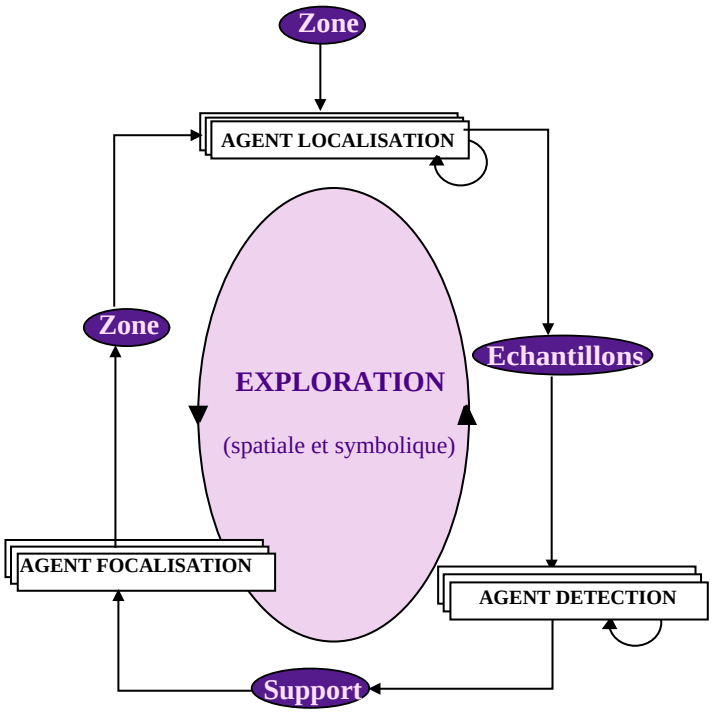


Fig. 4.5 – Exploration spatiale de l'image articulée autour de trois types d'agents : les agents localisation, les agents détection et les agents focalisation.

Détection. De la même façon suivant le nombre de focalisations à réaliser, l'agent Détection demandera la création d'agents focalisation, qui définiront les nouvelles zones d'intérêt, et l'objectif associé. Enfin, sur chaque zone d'intérêt, un agent Localisation sera créé et pourra rechercher les échantillons en fonction de ce nouvel objectif.

Ainsi, nous avons défini un processus itératif pour l'exploration spatiale de l'image qui effectue en permanence un basculement simultanément dans l'espace des buts et dans l'espace image. Le système cherche en fonction des règles de focalisations de nouvelles zones d'intérêt auxquelles sont associés de nouveaux objectifs, de façon à accumuler un maximum de connaissances. Ce sont précisément ces règles qui expriment le lien entre le "quoi" et le "où" autrement dit entre le but et la zone.

La fusion d'informations

Dans le cadre de notre application qui est la recherche de cibles dans l'image, des informations de natures différentes peuvent être considérées. Par exemple, dans une image infrarouge un véhicule peut être caractérisé par une zone chaude en mouvement. Il est toujours intéressant de tirer profit des informations disponibles en vérifiant les hypothèses courantes avec différentes sources d'informations telles que les segmentations issues de l'exploration spatiale et de la recherche de zones en mouvement. Cette boucle exploite les liens qui existent entre l'espace des buts et l'espace des méthodes.

De même que pour l'exploration spatiale, chaque étape de la fusion d'informations est prise en charge localement par un type d'agent. Deux nouveaux agents ont donc été spécifiés : les agents Décision et Mouvement.

L'agent Mouvement recherche le support des objets en mouvement alors que, l'agent Décision a pour but d'exploiter les deux supports extraits respectivement par les agents détection et mouvement. Il s'appuie sur les informations déjà collectées pour décider de la nouvelle action à effectuer.

- soit, l'information disponible est suffisante pour prendre une décision satisfaisante. Le processus de détection est alors terminé, le système bascule dans une autre phase d'analyse qui est la reconnaissance ;

- soit, les informations obtenues ne permettent pas de prendre une décision satisfaisante et dans ce cas un nouveau cycle d'analyse est demandé.

Cette boucle relative à la fusion d'informations permet au travers d'informations de natures différentes de vérifier ou de rejeter des hypothèses avec une variation de la satisfaction associée à chaque support. La satisfaction est évaluée par rapport un seuil spécifié dans la base de connaissance et appelé seuil de validation.

Un agent supplémentaire, nommé agent Concept permet de prendre en charge l'activation des deux boucles d'analyse. Une représentation complète du système est présentée figure 4.6. On distingue deux boucles d'analyse : l'exploration et la fusion d'informations. Elles correspondent respectivement aux enchaînements Localisation-Détection-Focalisation et Détection-Mouvement-Décision.

Nous avons décrit le cycle d'analyse intra-image par la mise en œuvre des boucles relatives à l'exploration et à la fusion d'informations. Nous allons désormais nous intéresser à l'analyse inter-image, c'est-à-dire à l'exploitation de la séquence d'images.

4.4.3 Le cycle d'analyse inter-image.

Trois problèmes sont à considérer dans le cycle d'analyse inter-image :
- la consommation de la séquence d'images ;
- l'exploration des images ;
- la mise à jour de la base de connaissances.

La consommation de la séquence d'images est réalisée par une boucle sur un intervalle de temps. Pour chaque pas temporel le processus intra-image se relance via un agent Concept qui prend en charge la mise en œuvre des boucles d'analyse intra-image.

La seconde problématique est de tirer profit du passé. Le système doit accumuler la connaissance pour ne pas refaire l'ensemble des traitements sur chaque image. Sous l'hypothèse caméra fixe, il faut considérer deux cas de figures suivant la cinématique associée au concept :

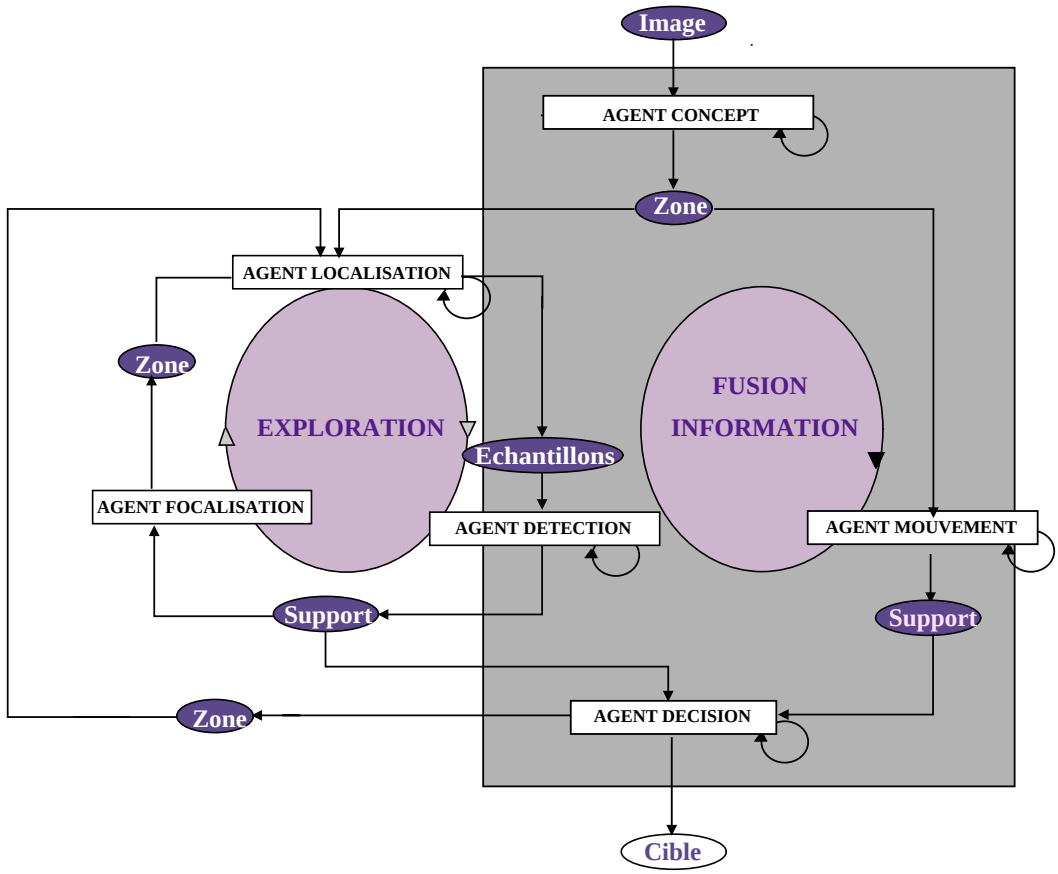


FIG. 4.6 – Le schéma de contrôle intra-image est comporte deux cycles d'analyse : un cycle d'exploration et un cycle de fusion d'informations.

- pour un concept fixe, nous avons choisi de conserver les échantillons satisfaisants obtenus par la population d'agents Localisation ; Ceci signifie que tous les échantillons extraits à l'instant t, associés à un concept fixe et ayant une confiance suffisante seront recopiés à l'instant t+1. Cette phase est contrôlée par un seuil de validation qui spécifie la confiance minimum autorisée. À partir des échantillons recopiés, une nouvelle analyse statistique est effectuée. Les mesures ainsi obtenues seront ensuite nécessaires à l'agent Détection pour réaliser sa propre segmentation.

- pour un concept en mouvement, il est judicieux d'effectuer une nouvelle analyse complète. Ainsi, aucune recopie n'est effectuée.

À chaque nouvelle image, le protocole de mise à jour de la base de connaissances est pris en charge par l'agent Concept. Cette mise à jour permet de réaliser une adaptation du système en fonction des informations accumulées et de faire évoluer pro-

gressivement la connaissance spécifiée *a priori*. Ainsi, les paramètres des méthodes peuvent notamment être modifiés, comme par exemple la plage de validation relative à l'intensité des pixels. Toujours pour se référer aux espaces de représentation c'est une adaptation opératoire dans l'espace du comment.

Cette description des cycles de contrôle du système permet de comprendre la philosophie globale de résolution du problème. Progressivement, nous allons détailler les différentes briques constituant ce système à savoir la base de connaissances, le modèle du monde et la gestion des incertitudes, et enfin la population d'agents.

4.5 Représentation des informations

Il existe un besoin constant de structurer les connaissances, de les regrouper pour caractériser une même entité, de représenter les liens de causalité ou les transitions. Nous allons donc, dans les paragraphes qui suivent, porter notre attention sur la représentation des différentes informations manipulées. Cette connaissance doit être structurée de façon à être accessible et disponible pour tous les agents du système. Nous distinguons deux types d'informations dans notre système :

- les connaissances descriptives et opératoires qui donnent une représentation *a priori* des différents concepts présents dans l'application ;
- les informations accumulées par les agents lors de l'analyse de la séquence d'images.

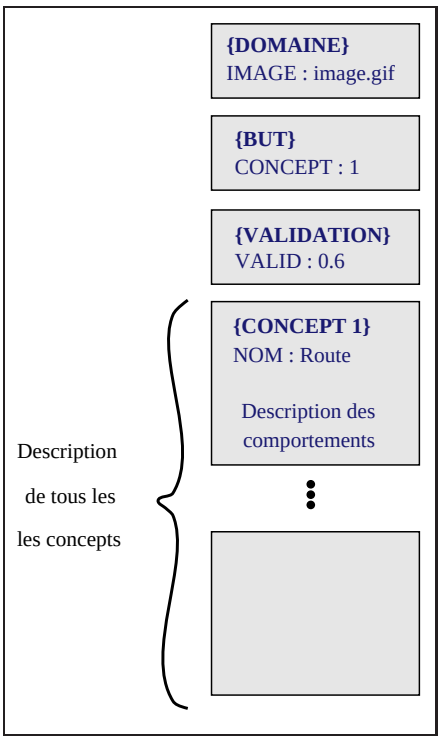
Cette distinction est importante puisque les modes de représentations associés à ces informations seront distincts. Et c'est précisément la composition de ces structures d'informations que nous allons détailler au travers de la base de connaissances et du modèle du monde.

4.5.1 La base de connaissances

La base de connaissances est spécifiée afin de fournir aux agents les informations nécessaires à leur mise en œuvre (figure 4.2). Elle définit le domaine d'analyse, l'objectif initial, le seuil de validation (utile dans le cycle de contôle inter-images) et donne une description opératoire des concepts [Touboul, 1998]. Compte tenu de l'étude présentée dans le paragraphe 2.2.1, cette base de connaissances intègre les modèles de scène qui diffèrent tant par leur nature (globale ou locale) que par les informations manipulées (géométriques, radiométriques, physiques ou encore symboliques).

Le principe :

La base de connaissances est structurée en quatre parties :



La première partie précise le domaine de travail, à savoir l'ensemble des images utilisées

La seconde partie définit l'objectif initial donné au système ;

La troisième partie donne le seuil de validation relative aux informations extraites ;

La quatrième partie décrit les concepts présents dans l'image en précisant leur nom et l'ensemble des méthodes utilisables par les agents, ainsi que les paramètres associés.

Le seuil de validation et la satisfaction seront utilisées notamment dans le cycle de contrôle inter-image, pour choisir les échantillons à recopier.

Les connaissances associées à chaque comportement sont décrites de la façon suivante :

```
COMPORTEMENT = [ COMP ] [ VALID ] [ CLAUSE ]
```

- `[ COMP ]` spécifie le comportement auquel est attachée la connaissance : différents comportements sont décrits il sont référencés par une lettre : 'L' pour la localisation, 'F' pour la focalisation, 'D' pour la détection et 'M' pour le mouvement ;
- `[ VALID ]` est un prédicat validant la mise en œuvre de cette connaissance ;
- `[ CLAUSE ]` décrit la méthode utilisée (transformée de Hough, niveaux de gris...) et ses contraintes.

Les méthodes utilisées sont spécifiées via des opérateurs de la bibliothèque de traitement d'images GAMMAG. [3]. Pour un même comportement, plusieurs clauses peuvent être combinées par un "ET logique". Un exemple est présenté figure 4.7. L'ensemble des méthodes et des paramètres sont décrits en annexe C.1 Les échantillons "Route" extraits par l'agent Localisation doivent impérativement vérifier trois conditions. La première est relative à la transformée de Hough (différentes approches possibles sont présentées en annexe C.2), la seconde à la valeur des niveaux de gris et enfin la troisième correspond à une condition géométrique sur la surface

[3] bibliothèque de traitement d'images interne au laboratoire DGA/CTA/GIP.

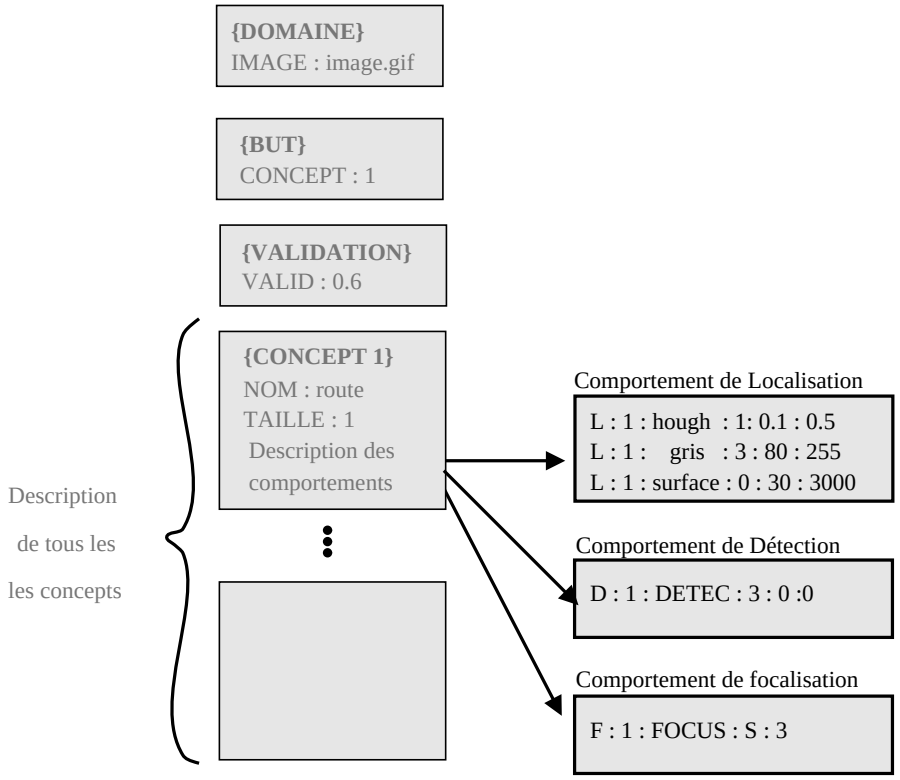


FIG. 4.7 – La base de connaissances décrit le domaine, l'objectif initial, le seuil de validation et donne une description opératoire de tous les concepts. Chaque comportement est caractérisé par un type, un prédicat de validation et un ensemble de méthodes spécifiant une condition et un intervalle de validation.

des échantillons. Suivant le schéma fonctionnel décrit sur la figure 4.6, à partir de chaque échantillon validé, une technique de détection est utilisée, pour trouver le support des objets. Ces supports seront exploités par l'agent Focalisation pour rechercher de nouvelles zones d'intérêt puisqu'un comportement de focalisation est spécifié. La syntaxe utilisée comporte deux paramètres : un but et un type de focalisation. Toujours dans cet exemple, la boucle de fusion d'informations ne sera pas exploitée, puisque aucune information sur la recherche du mouvement n'est précisée.

La spécification opératoire est souple. En effet le nombre de méthodes varie d'un agent à l'autre et d'un concept à l'autre. Une méthode peut être spécifiée ou non, validée ou non (avec le paramètre VALID) ou encore validée avec des paramètres différents suivant le concept considéré. La seule contrainte de fonctionnement qui existe, impose d'avoir au moins une méthode de localisation pour chaque concept décrit.

La structure de la base de connaissances permet d'établir directement deux liens dans l'espace de représentation (figure 4.8) :

- le lien entre “quoi” et “comment” : à chaque concept est associé un ensemble de méthodes et de paramètres qui sont ajustés localement et de manière autonome par l’agent.
- lien entre “quoi” et “ou” : la stratégie de focalisation établit le lien éventuelle entre l’espace image et l’espace des buts sous la forme d’une relation générique qui sera instanciée par les agents de localisation.

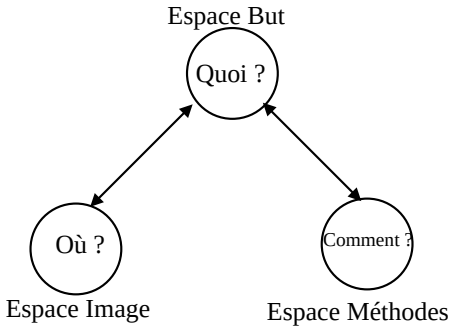


Fig. 4.8 – Base de connaissances : description des liens explicites entre les espaces de représentation

Chaque lien se décline à deux niveaux dans l’exploration du “où” et “comment”. Dans l’espace image, il y a une spécification de la méthode et un ajustement des paramètres. Dans l’espace des méthodes, il y a une spécification de la règle et une instanciation par l’agent Localisation.

L’ensemble des informations contenues dans la base de connaissances et mises à disposition des agents, est intimement lié au domaine d’application. Aussi, la première difficulté concerne la spécification de cette connaissance, puisque *a priori* on ne connaît pas le contexte applicatif. Notamment, suivant le type de mission (air-air, air-sol...) les éléments de scène seront différents et suivant la nature du capteur la perception pourra être modifiée.

Spécification de la base de connaissances

Une première solution possible est de procéder par apprentissage. On présente au système un ensemble de séquences tests, et il apprend progressivement à distinguer une route, d’un champ, ou d’un véhicule. Deux modes d’apprentissage s’opposent :

- un apprentissage “off-line”, est effectué sur une base test, au cours d’une phase d’entrainement, de préparation de mission ;
- un apprentissage “on-line” intégré au système de vision aggrège la connaissance au fur et à mesure des situations rencontrées. C’est notamment l’approche préconisée dans le système CITE [Dillion et Caelli, 1998].

La première technique nécessite un nombre d’images qui soit suffisamment important et représentatif des contextes étudiés. Ceci souligne une très grande difficulté

du problème. L'apprentissage "on-line" présente plusieurs avantages. L'information est accumulée de façon incrémentale suivant les situations rencontrées et autorise une adaptativité du système. Toutefois cette technique ne s'affranchit pas du problème de l'initialisation de la connaissance.

Pour pallier cette difficulté liée à l'initialisation, nous avons opté pour une acquisition interactive. Pour cela, nous utilisons une interface graphique (développée en TCL/TK). Le but est de sélectionner au mieux des zones d'intérêt représentatives d'un concept dans l'image. Sur chaque zone, la plupart des paramètres sont ensuite évalués de façon automatique. Ce procédé est itéré pour chaque concept étudié.

L'interface se décompose en trois modules distincts :

– Le premier module présente le domaine (figure 4.9).

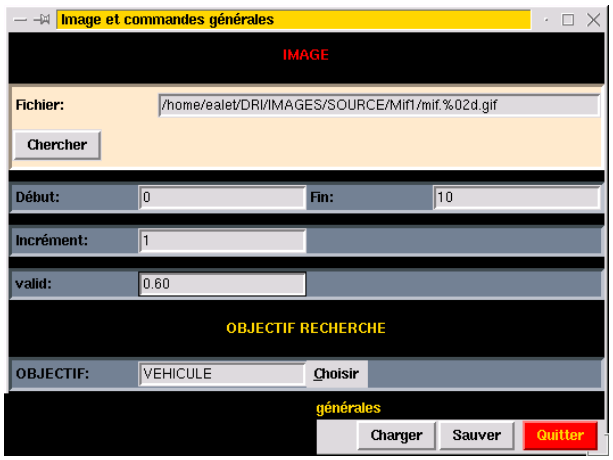


FIG. 4.9 – Premier module de l'interface graphique permettant de générer la base de connaissances. Il est défini par cinq champs : le nom de la séquence, les numéros de la première et de la dernière image, le pas d'incrémentation entre chaque image, le seuil de validation et enfin l'objectif initial.

Il permet de sélectionner la séquence d'images que l'on souhaite traiter en précisant les numéros de la première et de la dernière image analysée ainsi que le pas d'incrémentation. Par défaut les deux premiers paramètres référencent la séquence entière, et le pas d'incrémentation est fixé à 1. Enfin, l'opérateur choisit le seuil de validation, et l'objectif initial du système. Ce dernier paramètre doit impérativement appartenir à la liste des concepts définis.

– Le second module permet de visualiser une image de la séquence à partir de laquelle l'opérateur va constituer la connaissance *a priori*. Dans cette image, l'utilisateur peut choisir un rectangle d'intérêt et en extraire automatiquement une liste de paramètres statistiques. Un exemple est présenté sur la figure 4.10.

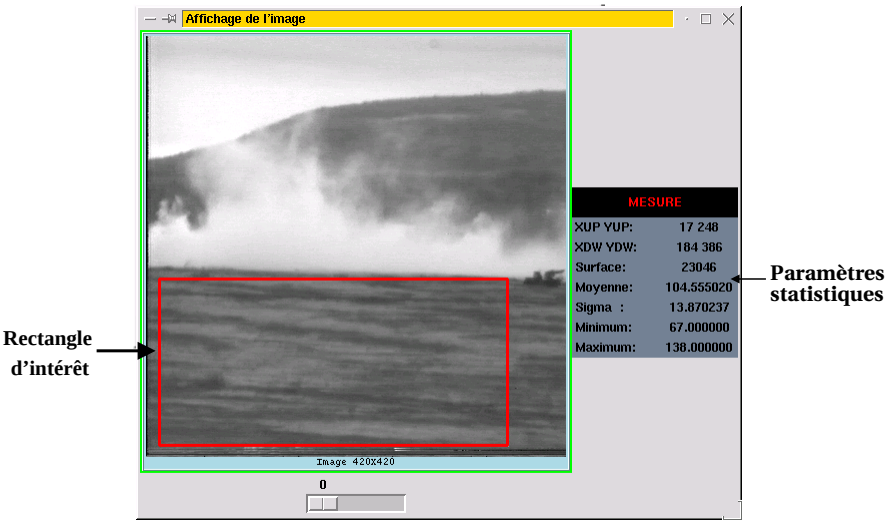


FIG. 4.10 – Le second module de l'interface graphique permet de visualiser l'image de référence et de visualiser les paramètres statistiques associés au rectangle d'intérêt défini en bas de l'image.

– Le dernier module permet véritablement de générer le fichier de connaissance. Pour chaque concept défini, l'opérateur sélectionne un ensemble de comportements à partir de l'interface présentée figure 4.11 (partie de gauche). Les ca-

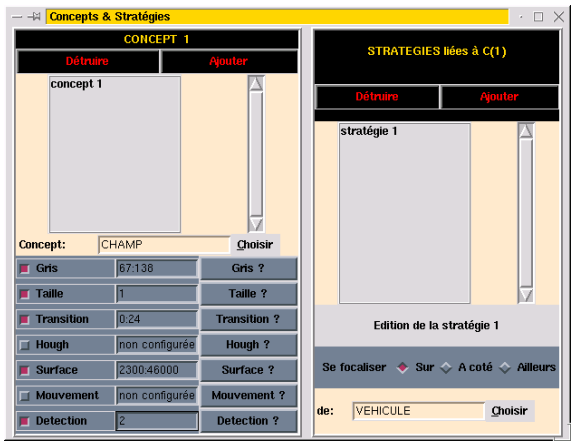


FIG. 4.11 – Troisième module de l'interface graphique permettant de générer la base de connaissances. La partie de gauche permet de sélectionner les comportements à mettre en œuvre, sachant que les clauses de validation sont ensuite générées automatiquement. Avec la partie droite de l'interface l'opérateur choisit les stratégies de focalisation.

ractéristiques obtenues sur la zone d'intérêt permettent de générer les clauses de validation de façon automatique pour chaque comportement. Ceci est possible pour la majorité des critères (à l'exception des paramètres utilisés pour la mise en œuvre de la transformée de Hough). Pour chaque concept, l'opérateur choisit également les stratégies de focalisation à savoir le type et l'objectif

associé (partie de droite sur la figure 4.11). Nous décrirons complètement ce problème dans la partie 4.6.4.

Le nombre de méthodes et le choix de ces méthodes est propre à chaque concept, mais il faut au moins qu'une méthode de localisation soit spécifiée.

Cette interface permet de générer un fichier de connaissances utilisé par les agents pour extraire les informations d'intérêt dans l'image.

À chaque nouvelle image, cette base de connaissances est réactualisée en fonction des informations collectées par les agents. Cette mise à jour est actuellement réalisée sur les intervalles de validation donnés pour chaque méthode. Une perspective est bien évidemment d'étendre cet apprentissage à la condition de validation de la méthode et au choix de cette méthode.

En résumé

La base de connaissances regroupe un ensemble de modèles de scène, décrits dans le paragraphe 2.2.1, qui différent suivant les informations manipulées :

- les modèles radiométriques (niveaux de gris et homogénéité) ;
- les modèles géométriques (surfaces, transformée de Hough) ;
- les modèles physiques (cinématique) ;
- les modèles structurels (stratégie de focalisation).

Ils diffèrent également par la nature des approches :

- Les modèles locaux regroupent notamment la caractérisation par niveaux de gris, par homogénéité, ou par transformée de Hough ;
- Les modèles globaux intègrent la cinématique, la surface des objets.

Cette base de connaissances est construite par le biais d'une interface graphique. À chaque concept sont associés une zone de l'image et un ensemble de paramètres qui permettront de spécifier automatiquement les différentes clauses. Une fois construite, elle est mise à disposition des agents du système et actualisée au cours du temps.

4.5.2 Le modèle du monde

L'information accumulée par les agents au cours du temps vient progressivement enrichir le modèle du monde pour donner une description de la scène. Cette structure de données est représentée sur la figure 4.12. Elle est composée de cinq sous structures :

- la structure Image est référencée par un identificateur. Elle contient l'image de référence et les images intermédiaires calculées (images Mouvement, Seg-

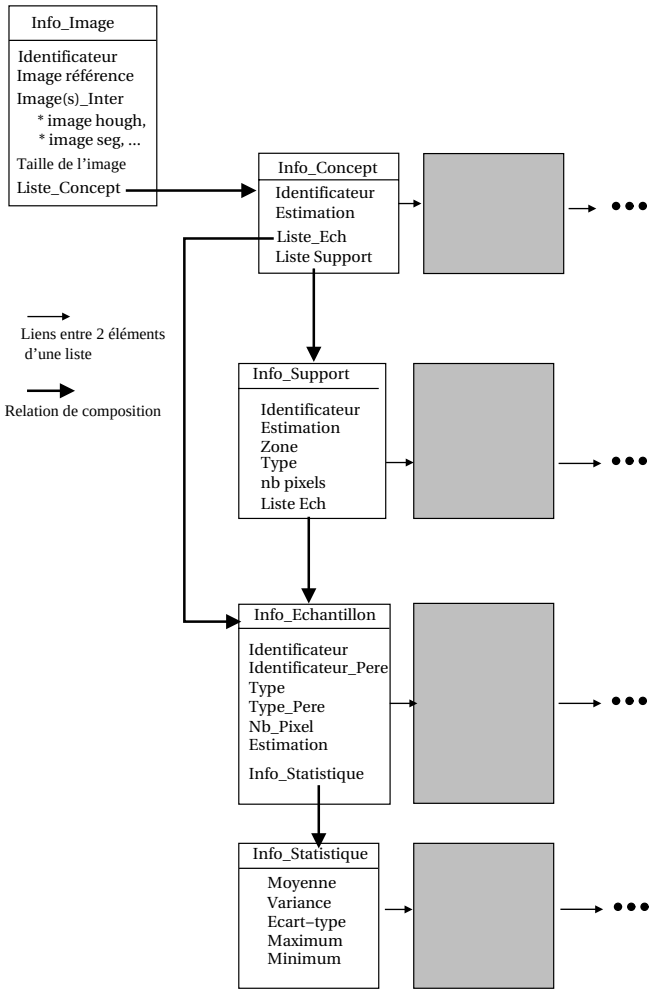


FIG. 4.12 – Structure du modèle du monde : pour chaque image, on peut accéder à la liste des concepts, des supports, des échantillons et des paramètres statistiques associés. Un concept est composé de plusieurs supports, eux-mêmes composés de plusieurs échantillons caractérisés par une liste d'informations statistiques.

mentation, Hough...). Elle dispose également de la liste des concepts qui ont été trouvés dans l'image.

– la structure Concept regroupe un paramètre d'identification, une mesure de confiance, et la liste des supports et des échantillons représentatifs du concept.

– la structure Support est caractérisée par un identificateur, un type et un nombre de pixels. Le rectangle englobant associé à chaque support est sauvegardé dans le paramètre "zone". Chaque support est également qualifié par une mesure de confiance qui représente la probabilité d'appartenance à un concept. Enfin, il contient la liste des échantillons qui lui ont donné naissance.

- la structure Échantillon est caractérisée elle aussi, par un identificateur, un type et une zone. De plus, elle contient les coordonnées de son centre de gravité, une liste d'informations statistiques et une mesure de confiance permettant d'évaluer la probabilité d'appartenance au concept de référence.

- la structure Statistique contient un ensemble d'informations statistiques telles que la moyenne, la variance, l'écart type, le minimum et le maximum. Ces informations peuvent être extraites des différentes images issues des résultats intermédiaires (image des niveaux de gris, image des transitions).

D'un point de vue pratique, ces structures sont représentées sous forme d'objet et gérées par des procédés de listes chaînées.

Cette description de la scène se construit de manière incrémentale. Dès qu'un agent obtient un résultat il le stocke. Il y a tout au long du processus d'analyse une mise à jour du modèle du monde. Les informations collectées sont caractérisées par une mesure de confiance qui permet d'évaluer le coefficient d'appartenance d'une entité à un concept donné. Ce paramètre évolue au cours du temps et se propage dans la structure d'informations par le biais de réseaux bayésiens.

La spécification complète d'un réseau bayésien suppose de :

- recenser toutes les variables aléatoires à considérer ;
- ordonner ces variables, à savoir distinguer les causes et les conséquences pour déterminer le graphe ;
- spécifier les probabilités conditionnelles à chaque noeud ;

Ainsi, il est possible de calculer les probabilités *a posteriori*, à partir des probabilités *a priori* en appliquant la règle de Bayes.

Ce formalisme nécessite, il est vrai la spécification complète du modèle probabiliste. Mais, les mécanismes de propagation associés à ces réseaux permettent de gérer les nouvelles observations du système, par instanciation de certaines variables. L'influence est propagée ensuite aux autres variables du réseau compte tenu des relations de dépendances "causes-conséquences" fixées.

Le calcul de la confiance associée à chaque support peut être donné sous forme d'un graphe de causalité. Il dépend de plusieurs sous réseaux (figure 4.13) :

- pour chaque échantillon appartenant au support, une confiance est calculée sur les attributs via un réseau bayésien (Ech1, ... EchN).
- chaque support dispose d'une filiation provenant des relations de focalisation. Par exemple, on renforcera facilement une hypothèse du type "véhicule sur route" alors que l'on minimisera la probabilité d'avoir "route dans ciel".
- enfin, l'information mouvement peut également renforcer une hypothèse telle que "véhicule en mouvement".

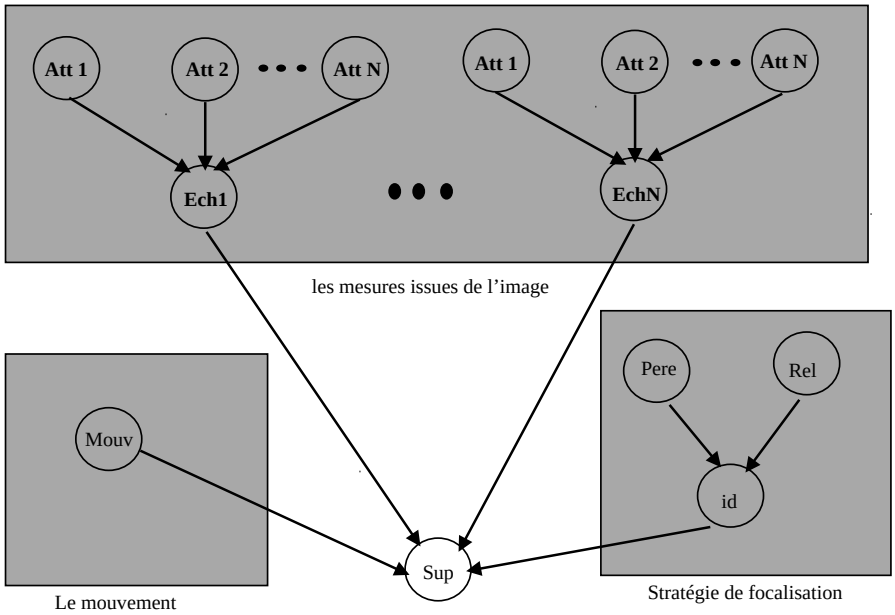


FIG. 4.13 – Description du réseau bayésien permettant le calcul de la confiance d'un support. Cette confiance dépend de la confiance obtenue pour chaque échantillon, des relations de focalisation, et du mouvement.

Suivant les informations calculées par les agents, les probabilités vont évoluer et être mises à jour. Nous détaillerons l'ensemble des sous-réseaux mis en œuvre dans le système lors de la description des comportements des agents.

4.6 Comportements des agents

Chaque agent évolue au cours du temps suivant les données dont il dispose. Pour réaliser son objectif, il utilise des comportements qui décrivent l'ensemble des actions possibles. Ces comportements lui sont spécifiques et lui permettent d'extraire directement ou indirectement des informations de l'image.

Nous distinguons, dans notre système, six types d'agents. Nous décrirons dans ce paragraphe cette population en décrivant notamment leurs différents comportements, les paramètres d'initialisation et les calculs d'incertitude.

4.6.1 L'agent concept

C'est le premier agent du système à travailler dans chaque nouvelle image. Cette caractéristique lui confère des impératifs. Il doit tirer profit des données déjà calculées et diriger la recherche de nouvelles informations dans l'image.

L'accumulation temporelle des informations

Le problème pour l'agent Concept est d'exploiter les informations pertinentes des images précédentes, de façon à ne pas refaire des traitements redondants. Pour cela, il recopie les échantillons qui répondent aux deux critères suivants :

- la confiance est suffisante ;
- le concept associé est fixe, c'est-à-dire que le champ mouvement n'est pas spécifié dans la base de connaissances.

La première condition est intuitive, il est évident que toutes les informations très largement entachées d'erreurs ne doivent pas être propagées dans le système. Le seuil de validation correspondant est spécifié *a priori* dans la base de connaissances et peut être modifié par l'opérateur.

La seconde condition résulte d'une hypothèse faite sur le contexte applicatif. En effet, nous sommes partis d'une hypothèse du type capteur fixe. Aussi, les concepts fixes se retrouvent d'une image à l'autre sans modification de position, même si les conditions de prise de vue, en général, peuvent subir des modifications telles que les variations de luminosité.

Enfin, les informations sélectionnées pour être dupliquées dans l'image suivante sont utilisées pour mettre à jour la base de connaissances. Connaissant les statistiques associées aux supports des objets, il est possible de restreindre ou d'élargir les intervalles de validation attachés aux connaissances opératoires. Cette mise à jour se fait par moyennage pondéré sur l'ensemble des supports recopiés. La pondération est fonction du nombre de pixels et de la confiance dans l'échantillon.

Activation des cycles de contrôle intra-image.

L'agent Concept contrôle la mise en œuvre des deux boucles d'analyse relatives à l'exploration intra-image à savoir l'exploration spatiale et la fusion d'informations. Cette activation est directement liée aux spécifications de la base de connaissances.

En résumé

Le schéma de principe relatif à l'agent Concept est donné sur la figure 4.14. L'objectif et la zone d'analyse constituent les données d'entrée de l'agent. Pour réaliser son objectif, l'agent Concept utilise certaines informations de la base de connaissances et du modèle du monde pour initialiser l'environnement courant et créer les agents nécessaires à l'exploration complète de l'image. On distingue les comportements de mise à jour du modèle du monde et de la base de connaissances ainsi que l'articulation des boucles d'analyse via l'activation des agents Localisation et Mouvement.

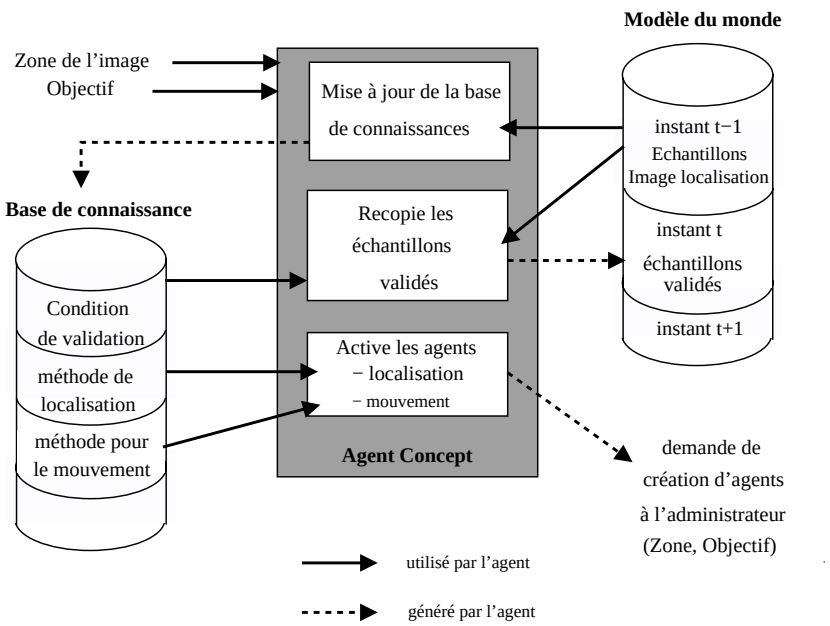


FIG. 4.14 – Schéma de principe de l'agent Concept. Trois comportements différents sont spécifiés : la mise à jour de la base de connaissances, la mise à jour du modèle du monde et enfin l'articulation des boucles d'analyse via l'activation de nouveaux agents.

4.6.2 L'agent localisation

L'agent Localisation met en œuvre deux comportements distincts : la localisation et l'estimation.

La localisation L :

Son rôle est d'identifier, dans une zone d'intérêt Z, un ensemble d'échantillons $\{E_i\}$, correspondant à un objectif courant O, soit :

$$\begin{array}{ll} \mathcal{Z} \text{ x } \mathcal{O} & \longrightarrow \mathcal{E} \\ (Z_I, O) & \xrightarrow{L} \{E_i\} \end{array}$$

Un échantillon est un ensemble de points vérifiant l'ensemble des clauses relatives au comportement de localisation et associées à l'objectif courant (voir le paragraphe 4.4.1). Celles-ci sont décrites dans la base de connaissances et spécifient les méthodes utilisées et leurs contraintes sous forme d'un intervalle de validation.

Pour valider un pixel de la zone d'intérêt, l'agent exploite l'information présente dans son voisinage. Un pixel n'est retenu que si un nombre suffisant N de voisins est validé.

La localisation est effectuée selon une approche locale, par déplacement d'une fenêtre de recherche de taille fixe. Cette taille spécifique à chaque concept est indiquée dans la base de connaissances. Le schéma de principe de la recherche des

échantillons est présenté sur la figure 4.15.

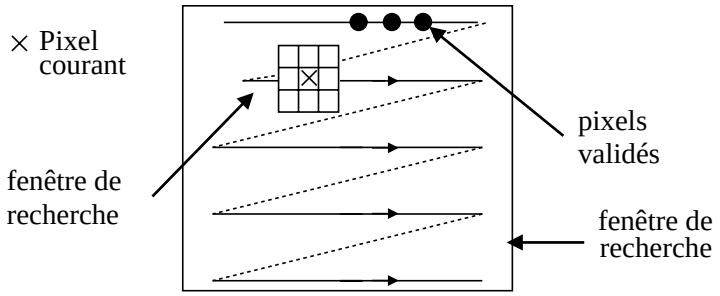


FIG. 4.15 – Principe de localisation des échantillons dans une zone d'intérêt. Un pixel est validé, lorsque chaque méthode disponible vérifie sa clause de validation dans une fenêtre de recherche.

Les méthodes mises en œuvre sont basées sur le niveau de gris des pixels, leur variance, ou à un niveau plus élevé sur la surface de régions et la longueur de segments.

Pour chaque méthode spécifiée dans la base de connaissances, l'image du résultat intermédiaire (homogénéité, transformée de Hough, ...) est générée et mise à disposition des autres agents dans le modèle du monde. L'agent Localisation procède par coopération confrontative (cf. paragraphe 3.4).

À la fin du processus, l'image de localisation est binaire. Elle distingue les pixels validés des autres. Pour obtenir l'information sous forme d'échantillons (régions élémentaires), l'agent Localisation procède alors à un étiquetage 4 ou 8-connexe sur la zone en question, suivi d'une analyse statistique classique. Chaque région est alors caractérisée par un identificateur, la zone d'intérêt, un nombre de pixels validés, une mesure de confiance E qualifiant l'appartenance à un concept et une liste de paramètres statistiques (attributs du premier ordre et second ordre, ...).

La confiance E :

À chaque échantillon extrait ($E_i \in \mathcal{E}$), nous associons une mesure de confiance évaluée comme une probabilité d'appartenance B_i ($B_i \in \mathcal{B}$, avec B lénsemble des croyances), telle que :

$$\begin{array}{ll} \mathcal{E} & \longrightarrow \mathcal{B} \\ E_i & \xrightarrow{E} B_i \end{array}$$

Plusieurs méthodes sont mises en œuvre pour définir les échantillons. Chacune fournit un résultat partiel et imprécis. Aussi, nous devons évaluer l'imprécision de ces différentes méthodes pour la propager ensuite au niveau de l'échantillon.

Chaque concept est décrit dans la base de connaissances par un ensemble de méthodes M définies sur un intervalle de validation $[MIN,MAX]$ présenté dans

le paragraphe 4.5.1. La répartition de la population des échantillons est supposée suivre une loi normale $\mathcal{N}(\mu_0, \widehat{\sigma_0})$ avec $\mu_0 = \frac{MIN+MAX}{2}$.

Soit X une variable aléatoire centrée et réduite qui vérifie $E(X) = 0$ et $V(X) = 1$. Pour $X > 0$, la probabilité de bonne détection est :

$$Pdetec = P(|X| < t)$$

Plus généralement, soit Y une variable aléatoire qui suit une loi normale de paramètre m et s, alors $\frac{(Y-m)}{s}$ suit une loi normale centrée réduite (figure : 4.16). Dans ce cas $E(Y) = m$ et $V(Y) = s^2$.

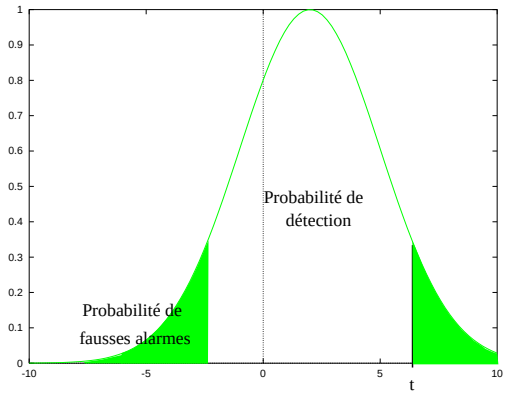


FIG. 4.16 – Distributions de probabilités associées au modèle de la base de connaissances

La probabilité de détection s'écrit :

$$Pdetec = P(|\frac{Y - \mu_0}{\widehat{\sigma_0}}| < t)$$

$$Pdetec = P(Y < (\mu_0 + t * \widehat{\sigma_0}))$$

Cette distribution de probabilités représente le modèle de la base de connaissances. Ainsi, l'agent Localisation va exploiter cette information pour estimer la probabilité de bonne détection sachant que la répartition de la population, issue de la mesure, pour chaque échantillon suit une loi normale $\mathcal{N}(\mu_{Im}, \sigma_{Im})$.

Une représentation graphique de ces deux distributions est proposée sur la figure 4.17.

Très schématiquement, nous voulons que la confiance caractérise l'écart entre ces deux courbes. Aussi, nous faisons l'hypothèse que la moyenne de l'échantillon correspond à l'une des bornes de l'intervalle de confiance relatif à la moyenne μ_0. Soit $\mu_{Im} > \mu_0$, on peut écrire l'équation suivante :

$$Pdetec = P(Y < (\mu_0 + t * \widehat{\sigma_0})) = P(Y < \mu_{Im})$$

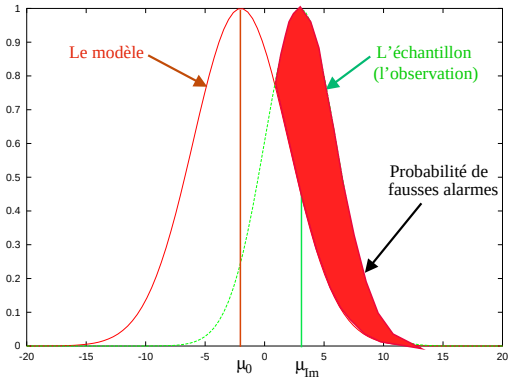


FIG. 4.17 – Distributions de probabilités associées aux informations

Par conséquent, puisque l'estimée est donné par $\widehat{\sigma_0} = \frac{\sigma_{Im}}{N}$, on a :

$$\mu_{Im} = \mu_0 + t * \frac{\sigma_{Im}}{\sqrt{n}}$$

Dans le cas général,

$$t = |\mu_{Im} - \mu_0| * \frac{\sqrt{n}}{\sigma_{Im}}$$

Et ainsi connaissant t, on détermine à l'aide des tables des fractiles, la probabilité de bonne détection.

Nous avons vu précédemment que la recherche d'échantillons dans une image s'appuyait sur la validation de plusieurs méthodes. À ce titre, l'incertitude associée à ces échantillons est égale à la moyenne de l'incertitude de chacune des méthodes. Ceci est décrit sous forme d'un réseau bayésien qui est présenté sur la figure 4.18.

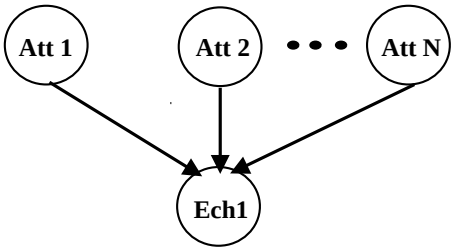


FIG. 4.18 – Calcul de la confiance d'un échantillon

Chaque noeud *Atti* représente l'incertitude sur la méthode i et prend deux états possibles représentés respectivement par V (vrai) et F (faux). On a :

$$P(Atti = V) = Prdi$$

$$P(Atti = F) = 1 - Prdi$$

La définition complète du réseau bayésien passe également par la description de la distribution de probabilités conditionnelles. Cette table est bien évidemment fonction du nombre de noeuds en entrée, mais elle se résume dans tous les cas de figure à une moyenne sur les mesures *Atti*.

Dans le cas simple où deux méthodes sont utilisées par l'agent localisation, comme par exemple le niveau de gris et l'homogénéité, le réseau est décrit de la façon suivante :

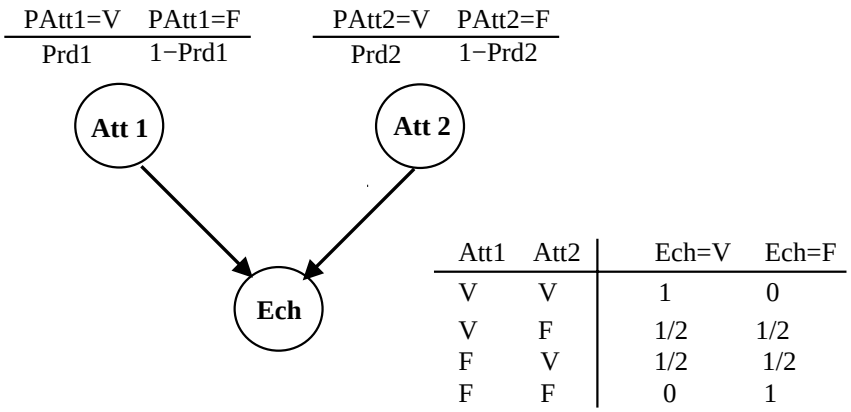


FIG. 4.19 – Le calcul de la confiance dans un échantillon est effectué via un réseau bayésien. Dans cet exemple, deux méthodes sont utilisées pour caractériser un échantillon.

Donc,

$$\begin{aligned} P(Ech = V) \quad &= P(Ech = V/Att1 = V, Att2 = V).P(Att1 = V).P(Att2 = V) \\ &+ P(Ech = V/Att1 = V, Att2 = F).P(Att1 = V).P(Att2 = F) \\ &+ P(Ech = V/Att1 = F, Att2 = V).P(Att1 = F).P(Att2 = V) \\ &+ P(Ech = V/Att1 = F, Att2 = F).P(Att1 = F).P(Att2 = F) \end{aligned}$$

L'outil utilisé pour la description des réseaux bayésiens est "the Bayes Net Toolbox" [4] (version 2.0) développé par K.P. Murphy à l'université de Berkeley. Une description est donnée en annexe A.

En résumé

Le schéma de principe de l'agent Localisation est présenté sur la figure 4.20. L'agent Localisation utilise une représentation locale des objets décrits dans la base de connaissances et les images de résultats intermédiaires issues de prétraitements (image de Hough, d'homogénéité, ...) stockées dans le modèle du monde pour valider des pixels de l'image de localisation.

[4] http ://www.cs.berkeley.edu/ murphyk/Bayes/bnt.html

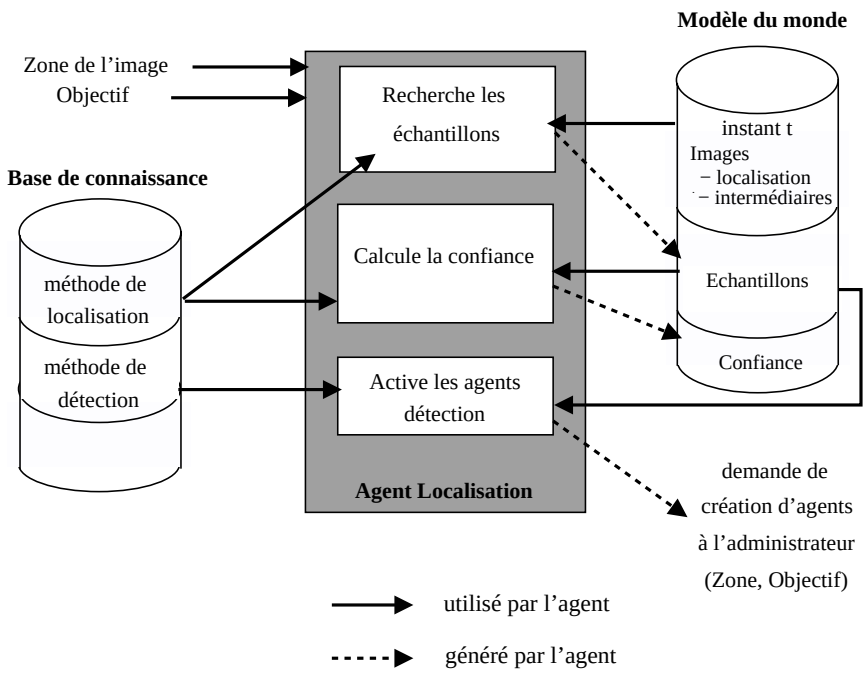


FIG. 4.20 – Schéma de principe de l'agent Localisation

Sur chaque région obtenue une analyse statistique est réalisée et une mesure de confiance est donnée via un réseau bayésien. Cette probabilité traduit l'appartenance d'un échantillon extrait à un concept donné connaissant l'incertitude liée à chacune des méthodes mises en œuvre. Dès qu'un échantillon est trouvé dans l'image, un agent Détection est activé afin de déterminer le support complet de l'objet.

4.6.3 L'agent détection

L'agent Détection dispose de deux comportements différents : la détection du support des objets et le calcul de la confiance associée.

La détection D : Son rôle est d'extraire le support d'un objet à partir d'un ou de plusieurs échantillons, soit :

$$\begin{array}{ll} \mathcal{E} & \longrightarrow \mathcal{S} \\ E_i & \xrightarrow{D} \{S_i\} \end{array}$$

La détection s'effectue selon une croissance de région à partir de l'échantillon extrait. Le point d'ancrage correspond au centre de gravité de l'échantillon et les critères de croissance sont définis par les informations statistiques calculées sur l'échantillon réalisant ainsi une adaptation locale. L'agent utilise la moyenne (moy), la variance (var) et un coefficient (coeff) spécifié dans la base de connaissances pour effectuer la croissance, suivant la règle :

$$(moy - coeff * var) < I < (moy + coeff * var)$$

Avec, I l'intensité du pixel courant.

L'agent détection s'adapte localement à son environnement. Il exploite l'information présente dans l'image pour ajuster ses paramètres.

Une fusion des étiquettes est nécessaire lorsque deux supports ont au moins un point de contact et référencent le même concept, suivant un principe de coopération augmentative. A la fin du processus le modèle du monde est mis à jour.

Si aucune méthode de détection n'est attachée au concept courant, le système assimilera un support à chaque échantillon extrait.

L'estimation E : Sa définition est similaire à celle présentée en 4.6.2. Le but est d'associer à chaque support S trouvé un coefficient d'appartenance B, soit :

$$\begin{array}{ll} \mathcal{S} & \longrightarrow \mathcal{B} \\ S_i & \xrightarrow{E} B_i \end{array}$$

La probabilité d'appartenance tient compte de trois paramètres :

- des confiances attachées aux échantillons qui ont donné naissance au support (cf. paragraphe 4.6.2) ;
- de la filiation qui a donné naissance au support. Cette information est donnée par un couple de données spécifiant l'identité du support père noté Sup_P et la stratégie de focalisation utilisée notée Rel. Par exemple, il est possible de chercher les véhicules sur les supports de type Route. Dans ce cas le couple décrivant la filiation est (route, sur) ;
- et enfin, de l'incertitude sur la segmentation elle-même.

Le réseau bayésien qui calcule la confiance pour un support est décrit sur la figure 4.21.

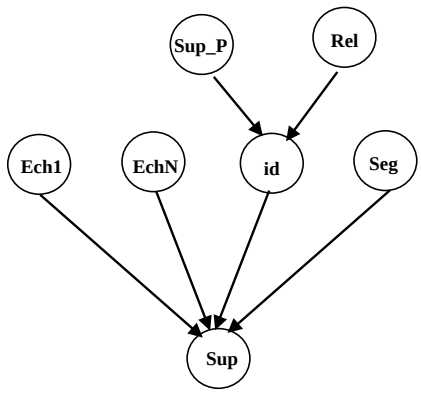


Fig. 4.21 – Réseau bayésien pour calculer la confiance dans un support

L'objectif est d'utiliser des connaissances de natures différentes pour renforcer ou non les hypothèses courantes. Ces connaissances regroupent les mesures statistiques

liées aux échantillons et au support, ainsi que des informations contextuelles *a priori*.

les trois principales hypothèses que l'on pourra utiliser sont :
- véhicule sur route
- bateau sur mer
- avion sur ciel

Les données contextuelles sont spécifiées dans la table de probabilités conditionnelles qui est indéxée par l'identité du support en question, l'identité du support père et la relation spatiale.

En résumé

L'agent Détection exploite les échantillons par l'intermédiaire du centre de gravité et des informations statistiques pour rechercher le support d'un objet. Le résultat de la segmentation est stocké dans le modèle du monde après une évaluation de l'incertitude. Un agent Détection est attaché à l'exploitation d'un seul échantillon. Le schéma de principe relatif à l'agent Détection est présenté sur la figure 4.22

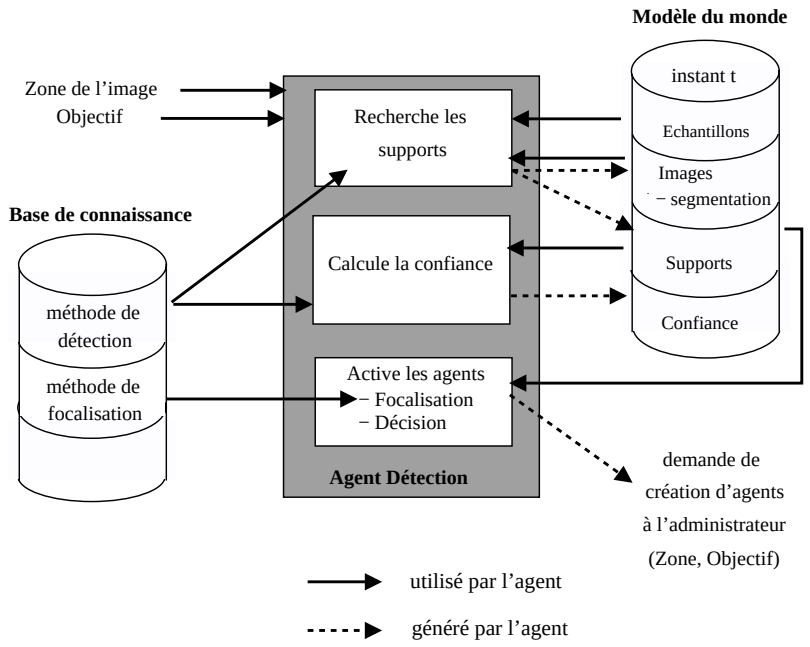


FIG. 4.22 – Schéma de principe de l'agent Détection. Il exploite les données relatives à un échantillon pour extraire le support d'un objet. Chaque résultat est évalué et fusionné si nécessaire avant d'être stocké dans le modèle du monde. Connaissant le support de l'objet en question, il est ensuite possible de définir de nouvelles zones d'intérêt via l'agent Focalisation et de prendre une décision sur la nature du support via l'agent Décision.

Dès que le support d'un objet potentiel est trouvé, un agent Décision est activé pour évaluer la pertinence de l'hypothèse en fonction des informations déjà collectées. Nous décrirons complètement le procédé utilisé dans le paragraphe 4.6.6. En

parallèle, tant qu'il existe des stratégies de focalisation associées à l'objectif courant qui n'ont pas été explorées, l'exploration spatiale est poursuivie par l'activation d'un agent de Focalisation.

4.6.4 L'agent focalisation

L'agent Focalisation permet de définir de nouvelles zones d'intérêt à partir des supports d'objet trouvés dans l'image (figure 4.23), soit :

$$\begin{array}{ll} \mathcal{S} & \longrightarrow \mathcal{Z} \text{ x } \mathcal{O} \\ S_i & \xrightarrow{F} \{(Z_i, O_i)\} \end{array}$$

Actuellement l'agent est doté de trois modes de focalisation. Il va se positionner soit sur le support (S), soit autour du support (A), soit ailleurs dans l'image (O) (figure 4.23).

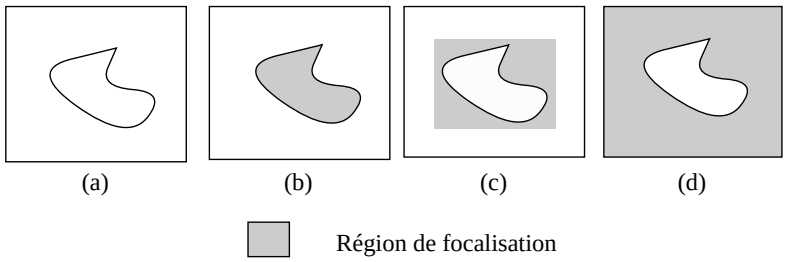


FIG. 4.23 – Principe de focalisation intra-image : (a) support de l'échantillon, (b) focalisation sur l'échantillon, (c) focalisation autour de l'échantillon et (d) focalisation ailleurs dans l'image.

Ces stratégies de focalisation, associées à chaque concept, sont spécifiées dans la base de connaissances. Elles sont décrites sous forme de couples de données associant le type de focalisation et le nouvel objectif. Par exemple pour rechercher les véhicules (2), sur un support déjà extraits, il faut spécifier la relation suivante :

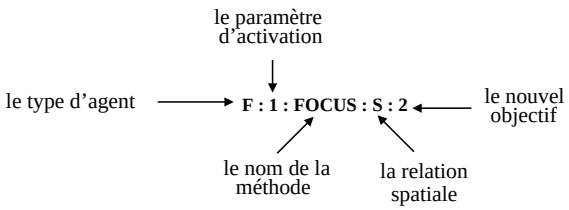


FIG. 4.24 – Exemple de stratégie de focalisation.

Les zones explorées sont sauvegardées dans une image appelée "image focalisation", et indexées par un concept. Il sera possible de se focaliser sur une région étiquetée si le concept recherché est différent de l'étiquette sauvegardée.

Ainsi à partir d'un support, l'agent focalisation fournit une nouvelle zone d'intérêt et l'objectif associé. Ces données seront les paramètres d'entrée d'un nouvel agent

Localisation. Cet agent Focalisation permet de progresser dans l'exploration spatiale et dans l'exploration symbolique des concepts suivant un principe de coopération intégrative (cf. paragraphe 3.4).

En résumé

Un agent Focalisation est activé lorsque deux conditions sont simultanément réalisées :
- l'agent Détection a trouvé un support relatif au concept C dans l'image ;
- au moins une stratégie de focalisation est spécifiée pour ce concept C.

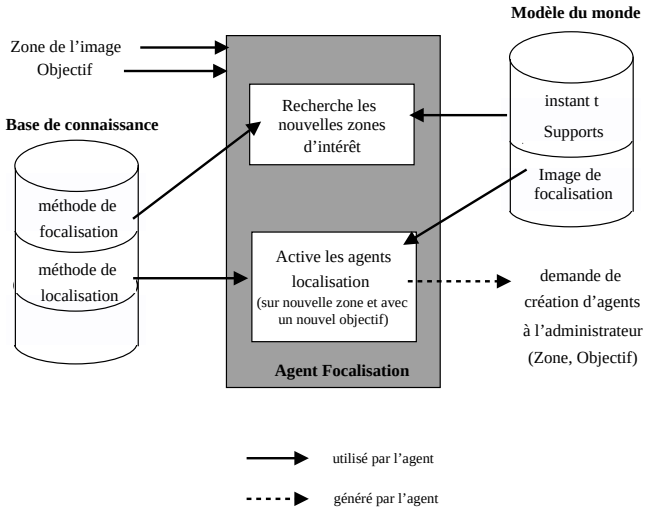


FIG. 4.25 – Schéma de principe de l'agent Focalisation. Il exploite le support d'un objet et la stratégie de focalisation décrite dans la base de connaissances pour définir une nouvelle zone d'intérêt et lui associer un objectif. L'analyse de l'image peut ainsi se poursuivre via un agent de localisation.

Chaque agent Focalisation exploite chaque support d'objet pour définir une nouvelle zone d'intérêt et donner le nouvel objectif de façon à poursuivre l'analyse de l'image avec la création d'un agent Localisation. Le schéma de principe est présenté sur la figure 4.25.

4.6.5 L'agent mouvement

L'agent Mouvement recherche le support des objets en mouvement et leur associe une mesure d'incertitude.

Recherche des supports d'objets en mouvement

L'algorithme utilisé a été proposé par Kornsprobst [Kornprobst *et al.*, 1997]. Il est très largement simplifié puisque la partie relative à la restauration du fond basé sur des méthodes variationnelles n'est absolument pas exploitée.

$$\begin{array}{ll} \mathcal{Z} \text{ x } \mathcal{O} & \longrightarrow \mathcal{S} \\ (Z_i, O_i) & \xrightarrow{M} \{S_i\} \end{array}$$

L'approche basée sur la contrainte du flot optique est formulée comme un problème d'optimisation. L'énergie est composée de :

- la contrainte du flot optique ;
- un terme de régularisation qui préserve les discontinuités et qui effectue un lissage dans les régions où le flot optique est homogène ;
- et enfin, un terme qui pénalise les vitesses lorsque aucune information texture n'est présente.

Cet algorithme, développé au sein du laboratoire CTA/GIP par L. Chanas, fournit l'image des champs de vitesses. Nous avons dit précédemment qu'il y avait autant d'agents Mouvement créés qu'il y avait de concepts en mouvement spécifiés. Bien évidemment, le calcul de l'image des champs de vitesse n'est quant à lui effectué qu'une seule fois. Ce résultat est d'ailleurs stocké dans le modèle du monde, et mis à la disposition de la population d'agents.

Un procédé d'étiquetage permet à l'agent d'extraire le support des objets en mouvement. À chaque support, suivant la même philosophie que précédemment, nous pouvons affecter une confiance dans la mesure.

Le comportement d'estimation

Pour calculer la mesure de confiance, l'agent Mouvement utilise le même procédé que celui décrit pour les méthodes de localisation. En fait, les champs de vitesse sont caractérisés dans la base de connaissances par un intervalle de validité qui est modélisé sous forme d'une gaussienne. L'objectif est donc de comparer le modèle avec la mesure obtenue (voir paragraphe 4.6.2).

En résumé

L'agent Mouvement est activé sur demande de l'agent Concept. Il peut être assimilé à un agent de veille qui très vite va demander une analyse pour prendre une décision, la plus rapide et la plus fiable possible. L'agent Mouvement va demander à l'administrateur la création d'un agent Décision, chaque fois qu'un support est trouvé. L'ensemble des caractéristiques de cet agent est décrit dans le schéma de principe présenté sur la figure 4.26.

4.6.6 L'agent Décision

L'agent Décision exploite différentes sources d'informations, telles que le mouvement ou l'information contextuelle, pour confirmer ou non des hypothèses. Il a un rôle décisionnel. Il peut arrêter l'exploration de l'image, ou demander une nouvelle

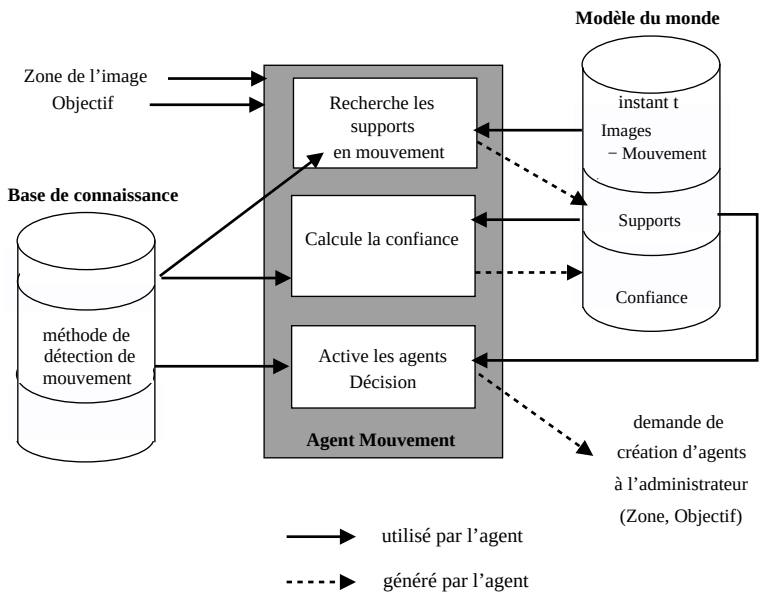


FIG. 4.26 – Schéma de principe de l'agent Mouvement. Il recherche le support des objets en mouvement. Pour chaque support, il demande la création d'un agent Décision, de façon à prendre rapidement une décision.

analyse.

Comme nous l'avons déjà signalé dans cet exposé, l'agent Décision peut être activé de deux manières différentes :

- soit, sur demande de l'agent Détection pour prendre une décision sur la présence d'une cible, ou d'un élément de la scène ;
- soit, sur demande de l'agent Mouvement pour prendre une décision également sur la présence d'une cible ou demander l'exploration de nouvelles zones d'intérêt dans l'image.

$$\begin{array}{ll} \mathcal{S}^n & \longrightarrow \mathcal{S} \\ \{S_i\} & \xrightarrow{D} S \end{array}$$

Lorsque pour un même support les hypothèses "mouvement" et "cible" sont validées, le système estime avoir détecté une cible et renforce la confiance associée. En revanche, si aucune de ces hypothèses n'est vérifiée, aucune cible n'est détectée. Si le support est en mouvement mais non segmenté, une analyse supplémentaire est nécessaire. Elle est réalisée par le cycle d'analyse localisation-détection-focalisation. En revanche, si un support cible est détecté, mais qu'aucune information mouvement n'est présente, l'hypothèse cible est diminuée et, suivant sa probabilité, elle est retenue ou non. Un schéma récapitulatif est présenté figure 4.27.

Entrées de l'agent Détection

Entrée de l'agent Mouvement

	CIBLE	NON CIBLE
MOUVEMENT	Probabilité augmentée	Poursuivre le recherche
NON MOUVEMENT	Probabilité diminuée	X

FIG. 4.27 – Comportement de l'agent Décision : les hypothèses émises en fonction de l'identité et de la cinématique du support fournies respectivement par les agents Détection et Mouvement conduisent à des actions différentes

En résumé

L'agent Décision exploite les informations obtenues pour évaluer le besoin de continuer ou non l'analyse. Il permet soit d'activer le processus de reconnaissance si une cible a été détectée, soit de poursuivre la phase de détection en activant la boucle relative à l'exploration spatiale. Le schéma de principe est présenté sur la figure 4.28.

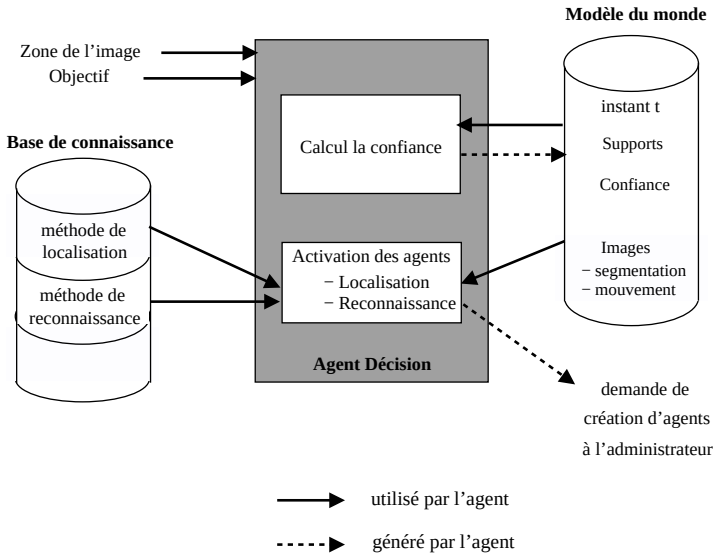


FIG. 4.28 – Schéma de principe de l'agent Décision

4.7 Conclusion

Notre architecture se décompose en trois structures : la plate-forme multi-agent, la base de connaissances et le modèle du monde.

– la plate-forme multi-agent gère la dynamique du système. Un agent est activé sur demande d'un autre agent. Il est détruit dès qu'il a fini son traitement. Les

agents se distinguent par leur rôle et les comportements qu'ils adoptent. Ils héritent des fonctionnalités de l'agent générique présenté dans le paragraphe 4.3.2 auquel sont ajoutées des spécificités suivant le rôle qui leur est attribué. Chaque agent exploite les structures d'informations pour réaliser son objectif et activer la poursuite de l'exploration de la séquence d'images.
- la base de connaissances fournit aux agents les informations nécessaires à leur mise en œuvre.
- le modèle du monde stocke les informations collectées par la population d'agents. Il s'enrichit au cours du temps et donne une description de plus en plus riche de la scène observée.

D'une manière générale, les agents sont situés dans un espace à plusieurs dimensions. Ils travaillent toujours dans une zone de l'image, avec un objectif et des méthodes qui leur sont propres.

La stratégie d'exploration et d'analyse n'est pas figée *a priori*. Elle dépend des décisions prises localement par les agents. Nous montrons dans le chapitre suivant, comment la modification de ces connaissances permet au système de développer des stratégies différentes.

Chapitre 5

Résultats expérimentaux

Sommaire

5.1 Introduction **111**
5.2 Quelques caractéristiques techniques **111**
5.3 Présentation du cadre d'application **111**
5.4 Exploration spatiale de l'image **114**
5.5 Exploitation d'une séquence d'images **127**
5.6 Conclusion **139**

Résumé

Après une description détaillée de notre architecture, nous proposons de valider expérimentalement notre approche, afin de mettre en valeur la structure de ce système, ses caractéristiques techniques et ses performances.

Nous allons considérer différents exemples pour mettre en valeur les boucles d'analyse liées à l'exploration spatiale et à la fusion d'informations.

Au cours de cette analyse, nous nous attacherons à montrer la distribution de notre système dans l'espace des buts, dans l'espace image et dans l'espace des traitements.

Nous porterons en permanence notre attention sur le lien mutuel qui existe entre les stratégies mises en œuvre et le calcul des probabilités.

Nous présenterons dans ce chapitre des résultats illustrant les principes de détection, de focalisation, d'incrémentalité et d'adaptativité, sur différents types d'images.

5.1 Introduction

Le système multi-agent mis en œuvre a été décrit complètement au cours des chapitres précédents. Nous avons présenté les différents modules qui le composent à savoir l'administrateur, les structures stockant les informations et la connaissance et enfin la population d'agents.

Pour reprendre la démarche de description, nous présenterons les résultats issus de la localisation spatiale, et l'amélioration des performances obtenues par la fusion et l'accumulation des informations.

Nous présenterons à chaque étape les liens entre les probabilités et les choix de stratégies et leur influence mutuelle.

Afin de valoriser l'architecture de notre système, ses caractéristiques techniques et ses performances, nous présenterons les résultats obtenus avec différents types d'images.

5.2 Quelques caractéristiques techniques

L'architecture multi-agent proposée est implantée suivant une philosophie distribuée de type multi-thread. La bibliothèque utilisée correspond à une implantation de la norme POSIX ("pthreads"). Elle fournit des primitives pour créer des processus légers et les synchroniser. Cette norme disponible sous l'environnement Linux et sur stations SUN autorise la portabilité du système.

Cette implantation est particulièrement intéressante dans le cadre d'applications multi-agent puisque l'idée est de distribuer les comportements et de les exécuter en parallèle.

Le système a été développé en C++. Là encore le choix d'un langage objet est adapté aux approches multi-agent. On assimile naturellement un agent à un objet. Le choix d'une programmation en C++ semblait également naturel pour interfacer le système avec la bibliothèque de traitement d'images GAMMAG, développée en C. Ce choix a permis également l'interfaçage avec la bibliothèque "Bayes Net Toolbox"(code matlab), pour la mise en œuvre des réseaux bayésiens.

L'expérimentation a été réalisée sur une station SUN, constituée de 32 processeurs, permettant de profiter pleinement du parallélisme. Un portage sous l'environnement Linux a également été validé.

5.3 Présentation du cadre d'application

Nous proposons d'expérimenter notre système sur une base d'images issues de capteurs infrarouges. Ces images sont caractérisées par des contextes très différents et des difficultés multiples. Elles sont présentées dans la figure suivante :

image n^o1

- capteur : infrarouge (polyphème)
- taille : 640*480
- nombre d'images : 1
- source : Aérospatiale
- nombre de concepts : 4
- nom des concepts : véhicule, route, végétation, champ
- difficulté : cibles très petites

image n^o2

- capteur : infrarouge
- taille : 640*480
- nombre d'images : 1
- source : Aérospatiale
- nombre de concepts : 4
- nom des concepts : véhicule, route, nuage, champ
- difficulté : petites cibles et présence de nuages

séquence n^o3

- capteur : inconnu
- taille : 420*420
- nombre d'images : 9
- source : DGA[1]
- nombre de concepts : 4
- nom des concepts : véhicule, champ, fumée, ciel
- difficulté : fumée en mouvement, véhicule en mouvement.

[1] Copyright Her Majesty the Queen in Right of Canada 1994, Defense Research Establishment Valcartier, Passive Electro-optical Systems Section, Electro-optics Division, Image Processing for Fire control Group.

Rappelons que le nombre et le nom des concepts sont choisis *a priori*, en revanche leurs paramètres sont obtenus par l'utilisation d'une interface graphique (décrite dans le chapitre précédent). L'ensemble des connaissances ainsi obtenues est stocké dans la base de connaissances et mise à disposition de la population d'agents.

Par exemple, la base de connaissances associée à l'image $n^0 1$ est présentée ci-après :

```
{DOMAINE}
RACINE :IMAGES/SOURCE/Vol/vol.%02d.gif
EXTENSION :gif :
DEBUT :0 :
NB :1 :
PAS :1 :
---------------------------------------
{BUT}
CONCEPT :4 :
---------------------------------------
{VALIDATION}
VAL :0.10
---------------------------------------
{CONCEPT 1}
NOM :VEGETATION :
TAILLE :1 :
L :1 :GRIS :2 :62 :81 :
L :1 :HOMOG :5 :0 :1 :
L :1 :SURFACE :0 :150 :10000 :

{CONCEPT 2}
NOM :CHAMP :
TAILLE :1 :
L :1 :GRIS :8 :72 :92 :
L :1 :HOMOG :8 :0 :10 :
L :1 :SURFACE :0 :1000 :30000 :
D :1 :DETECT :3 :0 :0

{CONCEPT 3}
NOM :VEHICULE :
TAILLE :1 :
L :1 :GRIS :1 :114 :255 :
L :1 :HOMOG :3 :1 :10 :
L :1 :SURFACE :0 :10 :500 :

{CONCEPT 4}
NOM :ROUTE :
TAILLE :1 :
L :1 :HOUGH :1 :0.1 :0.5 :
L :1 :GRIS :4 :80 :255 :
L :1 :SURFACE :1 :250 :4000 :
D :1 :DETECT :2 :0 :0
F :1 :FOCUS :S :3 :
```

Ce fichier est composé de 4 parties (cf. paragraphe 4.5.1) qui sont :

– la description du domaine qui donne le nom de l'image, le format de l'image, le

numéro de la première image, le nombre d'images et enfin le pas d'incrémentation entre les images. Ce dernier paramètre est utile, si l'on veut traiter uniquement certaines images de la séquence (par exemple, une sur deux).

- la spécification du but initial fixe le premier objectif du système. Rappelons que cet objectif correspond à un élément de la liste des concepts. Dans cet exemple, l'objectif initial est 4 et correspond au concept Route.

- la condition de validation est utile pour la gestion de la séquence d'images. En effet, nous avons déjà précisé que pour chaque nouvelle image traitée, il y a une mise à jour de la base de connaissances et une recopie des échantillons. Les informations utilisées lors de cette mise à jour doivent toutefois avoir une confiance suffisante qui doit être supérieure à ce seuil de validation.

- la description des concepts spécifie les différentes méthodes nécessaires aux agents et leurs paramètres (cf annexe C.1). On distingue les méthodes de localisation, de détection et de focalisation, caractérisées respectivement par les marqueurs 'L', 'D' et 'F'. Les méthodes de localisation sont basées sur des critères d'intensité (GRIS), d'homogénéité (HOMOG), et de surface (SURFACE). Les méthodes de détection sont basées sur la croissance de régions (DETEC). Et enfin les méthodes de focalisation (FOCUS) recherchent les nouvelles zones d'intérêt, à partir de l'information courante, suivant trois stratégies différentes : sur ('S'), à côté (A), ou ailleurs (O) dans l'image.

Dans ce chapitre, nous proposons d'analyser les réponses du système en fonction des domaines, des buts, des conditions de validation et des stratégies de focalisation.

5.4 Exploration spatiale de l'image

Étant donné un but initial, l'objectif final est de poursuivre l'exploration de la séquence tant qu'il est utile de le faire. Cette exploration est guidée par des règles de focalisation et un seuil de satisfaction. Un concept but pourra ainsi être atteint par une stratégie "directe" s'il est le premier concept spécifié, ou par une stratégie "indirecte" s'il est spécifié au sein de règles de focalisation. Suivant les stratégies utilisées, les réponses du système sont différentes. Nous illustrerons ces propos avec trois expérimentations différentes :

- avec l'analyse directe, le système cherche directement les véhicules dans l'image, sans s'aider des éléments de contexte.

- avec l'analyse indirecte :
  • le système utilise une seule stratégie de focalisation : chercher les véhicules sur les routes.
  • le système utilise deux stratégies de focalisation : chercher les véhicules sur

et à côté des routes.

L'ensemble des expérimentations sera effectué sur les images présentées dans le paragraphe précédent 5.3.

5.4.1 Analyse directe

La première expérience est très simple. Elle consiste à trouver les véhicules dans l'image en utilisant exclusivement les données de la base de connaissances attachées à ce concept. L'objectif initial et l'objectif final sont confondus. Le fichier de connaissances utilisé est le suivant :

```
{DOMAINE}
RACINE :IMAGES/SOURCE/Vol/vol.%02d.gif
EXTENSION :gif :
DEBUT :0 :
NB :1 :
PAS :1 :
---------------------------------------
{BUT}
CONCEPT :3 :
---------------------------------------
{VALIDATION}
VAL :0.10
---------------------------------------
{CONCEPT 1}
NOM :VEGETATION :
TAILLE :1 :
L :1 :GRIS :2 :62 :81 :
L :1 :HOMOG :5 :0 :1 :
L :1 :SURFACE :0 :150 :10000 :

{CONCEPT 2}
NOM :CHAMP :
TAILLE :1 :
L :1 :GRIS :8 :72 :92 :
L :1 :HOMOG :8 :0 :10 :
L :1 :SURFACE :0 :1000 :30000 :
D :1 :DETECT :3 :0 :0

{CONCEPT 3}
NOM :VEHICULE :
TAILLE :1 :
L :1 :GRIS :1 :114 :255 :
L :1 :HOMOG :3 :1 :10 :
L :1 :SURFACE :0 :10 :100 :
D :1 :DETECT :1 :0 :0

{CONCEPT 4}
NOM :ROUTE :
TAILLE :1 :
L :1 :HOUGH :1 :0.1 :0.5 :
L :1 :GRIS :4 :80 :255 :
L :1 :SURFACE :1 :250 :4000 :
D :1 :DETECT :2 :0 :0
```

L'image analysée est l'image $n^0 1$ présentée dans le paragraphe 5.3. Dans cette

image, quatre concepts sont définis mais il n'y a aucune relation de focalisation entre ces éléments. Seuls les attributs relatifs au concept véhicule seront exploités. Ils sont représentés en gras dans la base de connaissances.

Les modèles utilisés pour trouver les véhicules dans l'image, sont statistiques (niveaux de gris et homogénéité) et géométriques (surface). Ils sont utilisés en phase de localisation.

La population d'agents mise en œuvre, dans cet exemple, regroupe quatre types d'agents :

- un agent Concept demande l'activation de l'exploration spatiale dans toute l'image par l'intermédiaire d'un agent de Localisation. Aucune information mouvement n'est exploitée puisque aucun comportement Mouvement n'est spécifié.

- un agent Localisation recherche dans l'image l'ensemble des pixels vérifiant les clauses spécifiées dans la base de connaissances pour l'objectif véhicule. Après une phase d'étiquetage, il fournit des régions élémentaires appelées échantillons, auxquelles sont associées des informations statistiques. Dès qu'un échantillon est identifié et pour chacun d'eux, l'agent Localisation demande la création d'un agent Détection ;

- chaque agent Détection exploite un échantillon pour extraire un support d'objet. Aucune stratégie de focalisation n'est spécifiée, aussi l'exploration spatiale s'achève ;

- l'agent Décision ne dispose pas d'informations complémentaires de type mouvement par exemple. Il transmet donc les données obtenues sans demander d'analyse complémentaire.

Tous ces agents ne sont pas présents au même instant dans l'image. Dès qu'un agent a terminé sa tâche, il demande à l'administrateur de supprimer sa fiche d'information avant de se tuer. L'enchainement des étapes est présenté sur la figure 5.1.

Le résultat de segmentation obtenu est présenté sur la figure 5.2. Cette image de segmentation présente une visualisation des huit supports véhicules trouvés dans l'image. À chaque support est associée une probabilité d'appartenance au concept véhicule. Celle-ci est représentée par une couleur.

Les véhicules sont détectés avec une probabilité d'appartenance supérieure à 67%. En revanche, ces informations sont noyées dans une grande quantité de fausses alarmes qui ont des probabilités comprises entre 63 et 66%. Une illustration du modèle du monde obtenu est présenté en annexe D

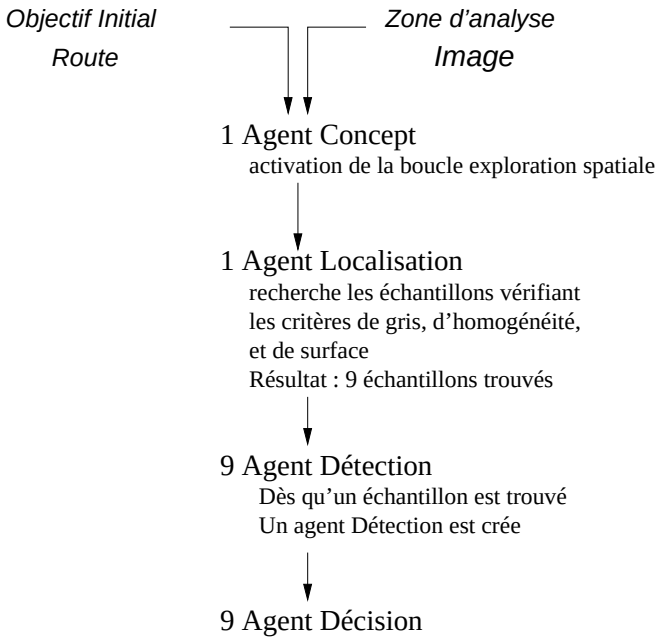


FIG. 5.1 – Enchainement des étapes pour l'analyse directe.

Le graphe représentant la population d'agents créés par le système pour détecter les véhicules est présenté sur la figure 5.3. Un lien entre deux agents représente une relation de type père-fils. Le père peut demander la création d'un fils pour continuer

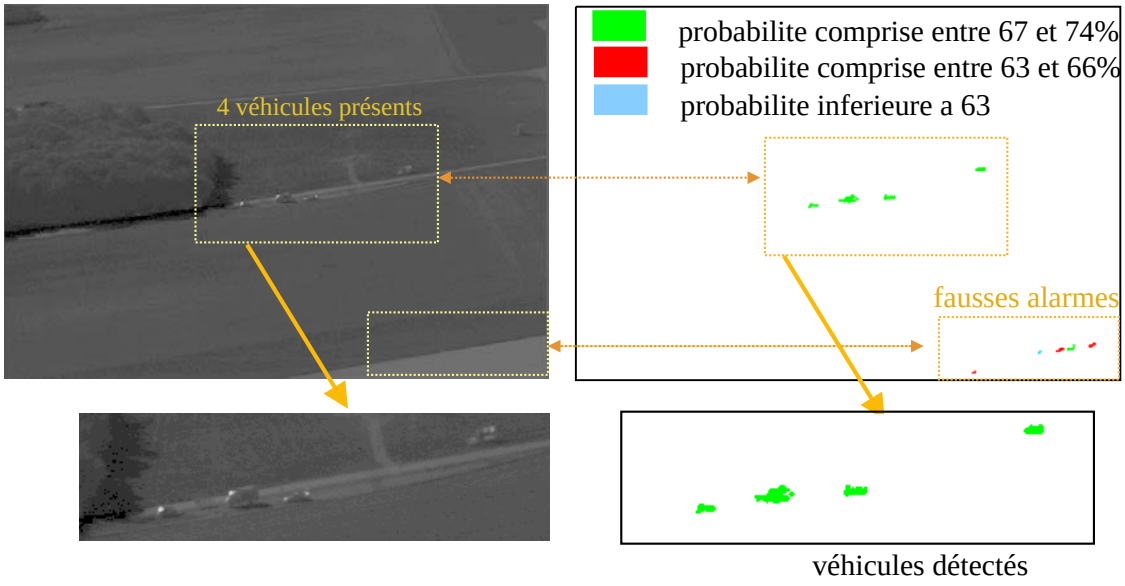


FIG. 5.2 – Segmentation des supports véhicule. À gauche, l'image initiale est présentée. Elle montre la présence de quatre véhicules. À droite, on a l'image de segmentation obtenue. La couleur associée à chaque support représente une probabilité d'appartenance au concept.

l'analyse.

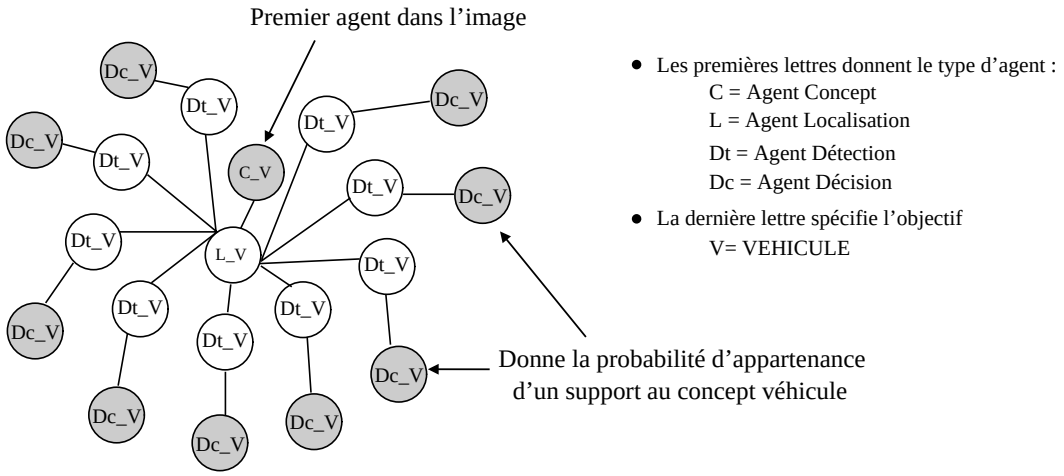


FIG. 5.3 – Graphe d'agents pour détecter les véhicules : il est basé sur la mise en œuvre de quatre types d'agents : les agents Concept, Localisation, Détection et Décision. Tous ces agents ont le même objectif à savoir le concept véhicule.

Chaque agent décrit est ancré dans un espace à trois dimensions :

- l'espace image : un agent travaille dans une région de l'image. Les agents Concept et Localisation exploitent toute l'image. Les agents Détection et Décision ne s'intéressent qu'à certaines régions de l'image qui sont définies en fonction du résultat fournit par l'agent Localisation ;

- l'espace des concepts : tous les agents ont le même objectif : véhicule ;

- l'espace des méthodes : chaque type d'agent met en œuvre des méthodes très différentes décrites en annexe C.1.

Au travers de cette population d'agents, une décentralisation apparaît très tôt dans le système. Le contrôle est pris en charge localement par chaque agent. D'autre part, il y a une distribution des étapes de traitements qui permet de répartir la charge de calculs sur des processus différents. Cependant, l'étape de localisation reste très coûteuse en temps de calcul, et notamment lorsque celle-ci s'effectue sur toute l'image. Le tableau ci-dessous donne les ordres de grandeurs des temps obtenus pour l'exemple que nous venons de dérouler.

Temps réel	19.2s
Temps utilisateur	16,2s
Temps système	2s

Cette lourdeur de l'étape de localisation est non négligeable. D'où l'intérêt de la

focalisation, qui fournira au système des zones d'analyse plus restreinte et diminuera le temps de calcul. En effet, la localisation n'est effectuée qu'une seule fois sur la totalité de l'image, même dans le cas d'une séquence d'images et uniquement pour le premier concept recherché.

Toujours dans le cadre de notre exemple, la répartition des temps de calculs affectés aux différents agents est présentée sur la figure 5.4.

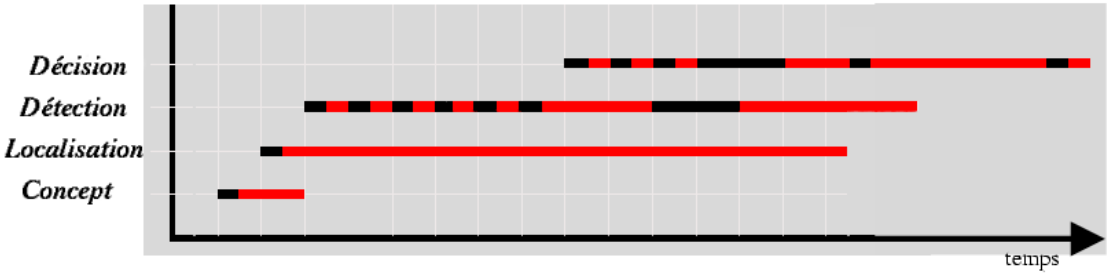


FIG. 5.4 – Répartition des temps de calculs entre les différents agents : en noir est représentée l'activation d'un agent, et en rouge son exécution.

On constate que l'on attend pas la fin de la localisation pour activer les agents de détection. Toutefois dans cet exemple, la parallélisation est relativement limitée puisqu'elle n'est présente qu'aux étapes de détection et de décision. Il est vrai que la parallélisation des traitements au sein même des agents est une étape qui s'avérera indispensable dans l'avenir. Pour l'instant, celle-ci n'a pas été réalisée pour des problèmes de portabilité.

Le résultat de segmentation obtenu est quant à lui très moyen. Certes les véhicules sont détectés malgré le peu d'informations dont on dispose, mais le taux de fausses alarmes est important (44%). Nous avons renouvelé l'expérience sur différentes images. Les résultats obtenus sont relativement similaires. On oscille systématiquement entre une sur-détection et une sous-détection. L'expérience sur l'image $n^0 2$ illustre le problème de la sous-détection. Le résultat de la segmentation est donné sur la figure 5.5. Des zones de l'image (pointée par des flèches rouges) pouvant représenter une menace ne sont pas considérées.

La quantité d'informations relatives à la cible est très faible et celle-ci est noyée dans un environnement complexe. Pour répondre à cette difficulté, et contraindre la recherche, une solution est de tirer profit de l'information contextuelle présente dans l'image.

5.4.2 Apport du contexte

Pour contraindre le domaine de recherche des informations d'intérêt, nous avons introduit des modèles traduisant les relations contextuelles entre les différents con-

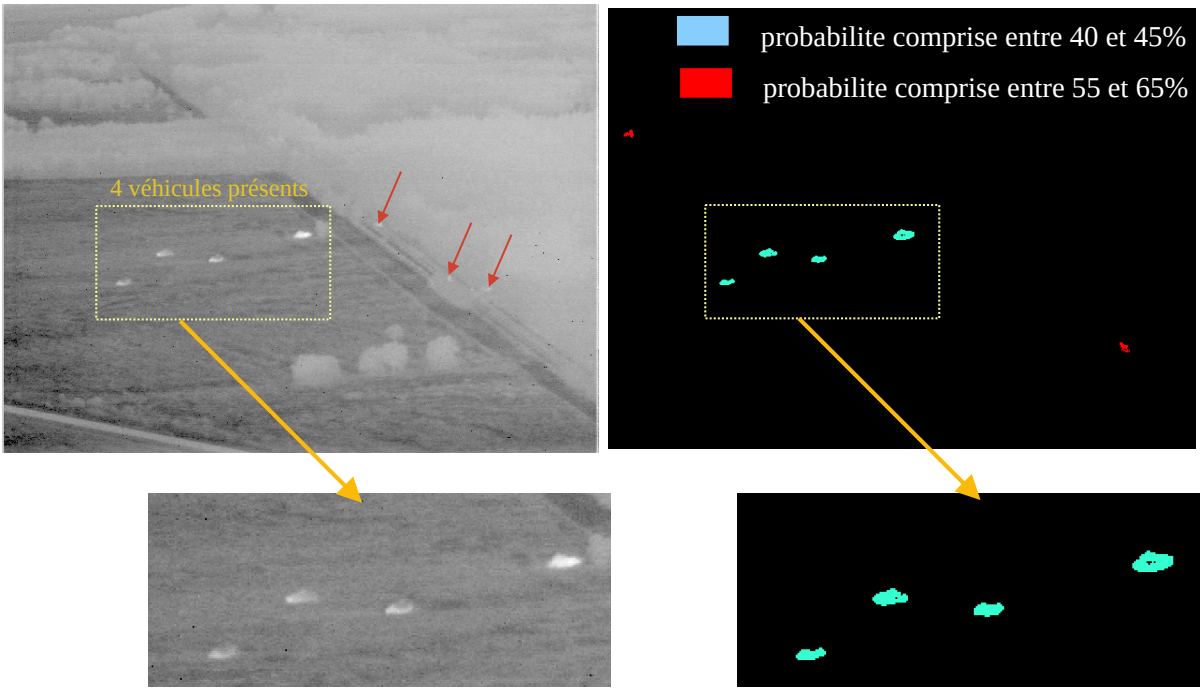


Fig. 5.5 – Autre segmentation des supports véhicule. À gauche, l'image initiale est présentée. Elle montre la présence de quatre véhicules. À droite, on a l'image de segmentation obtenue. La couleur associée à chaque support représente une probabilité d'appartenance au concept.

cepts présents dans la scène. Ces relations, attachées à un concept, sont exprimées sous formes de stratégies de focalisation. Elles sont spécifiées dans la base de connaissances.

Rappelons que, dans notre système, trois stratégies de focalisation sont définies. Il est possible de se focaliser : "sur" un support, "à côté" d'un support, ou encore "ailleurs" dans l'image. Nous présentons dans ce paragraphe les résultats d'expérimentation obtenus avec une seule, puis avec plusieurs stratégies de focalisation.

Avec une règle de focalisation

La première idée est de rechercher les grandes structures de l'image où le système a de fortes chances de trouver les véhicules. Reprenons l'exemple de l'image $n^0 1$ présentée dans le paragraphe 5.3. On choisit comme objectif initial les routes, sachant que la stratégie de focalisation associée sera de rechercher les véhicules sur les routes. Le fichier associé à cet exemple et représentant la base de connaissances, est présenté ci-dessous :

```
{DOMAINE}
RACINE :IMAGES/SOURCE/Vol/vol.%02d.gif
EXTENSION :gif :
DEBUT :0 :
NB :1 :
PAS :1 :
---------------------------------------
{BUT}
CONCEPT :4 :
---------------------------------------
{VALIDATION}
VAL :0.10
---------------------------------------
{CONCEPT 1}
NOM :VEGETATION :
TAILLE :1 :
L :1 :GRIS :2 :62 :81 :
L :1 :HOMOG :5 :0 :1 :
L :1 :SURFACE :0 :150 :10000 :

{CONCEPT 2}
NOM :CHAMP :
TAILLE :1 :
L :1 :GRIS :8 :72 :92 :
L :1 :HOMOG :8 :0 :10 :
L :1 :SURFACE :0 :1000 :30000 :
D :1 :DETECT :3 :0 :0

{CONCEPT 3}
NOM :VEHICULE :
TAILLE :1 :
L :1 :GRIS :1 :114 :255 :
L :1 :HOMOG :3 :1 :10 :
L :1 :SURFACE :0 :10 :100 :
D :1 :DETECT :1 :0 :0

{CONCEPT 4}
NOM :ROUTE :
TAILLE :1 :
L :1 :HOUGH :1 :0.1 :0.5 :
L :1 :GRIS :4 :80 :255 :
L :1 :SURFACE :1 :250 :4000 :
D :1 :DETECT :2 :0 :0
F :1 :FOCUS :S :3
```

Dans cet exemple, l'objectif initial est de rechercher les routes dans l'image. Pour cela, trois méthodes de localisation sont utilisées : utilisation de la transformée de Hough, de l'intensité des pixels et de la surface des régions. Ensuite, une méthode de détection permet d'extraire le support complet des concepts routes. Il est important de souligner, enfin, la clause de focalisation. Elle précise de rechercher à partir des supports route obtenus et avec une stratégie du type "sur", le concept véhicule numéroté 3.

Même si la recherche des routes peut paraître coûteuse en temps de calcul, nous allons montrer l'intérêt avec les performances obtenues.

La population d'agents est toujours ancrée dans un espace à trois dimensions. Dans cet exemple, on distingue cinq types d'agents : les agents Concept, Localisa-

tion, Détection, Focalisation et Décision. Il y a un agent supplémentaire par rapport à l'exemple précédent qui est l'agent Focalisation. Il va permettre le changement d'objectif au cours de l'analyse et le changement de zones d'intérêt.

Chaque agent travaille dans une région de l'image avec un objectif qui lui est propre. Dans le cas présent, l'espace des buts B contient deux éléments : B={Véhicule, Route}. Afin de bien souligner cet ancrage dans l'espace des buts, nous associerons désormais à chaque type d'agent son objectif. Nous parlerons par exemple d'agent du type localisation véhicule. Nous pouvons encore détailler un peu plus cette population en décrivant chronologiquement les agents présents dans le système et en reprenant cette notation.

- l'agent Concept Route recherche la zone d'intérêt initiale qui comprend par défaut, toute l'image ;

- l'agent Localisation Route recherche les échantillons vérifiant le concept route. À partir des échantillons trouvés, le système va procéder à une détection des supports.

- les agents Détection Route donnent les supports des routes et autorisent la poursuite de l'exploration spatiale via des agents de type Focalisation Route ;

- chaque agent Focalisation Route définit une nouvelle zone d'intérêt et le nouvel objectif (véhicule) à fournir à un agent de Localisation. Puisque la stratégie est du type "sur", la zone d'intérêt correspond aux supports route extraits. Les zones qui vont être analysées avec l'objectif véhicule sont représentées en bleu sur la figure 5.6. Il y a trois supports route, donc il y a trois agents Focalisation

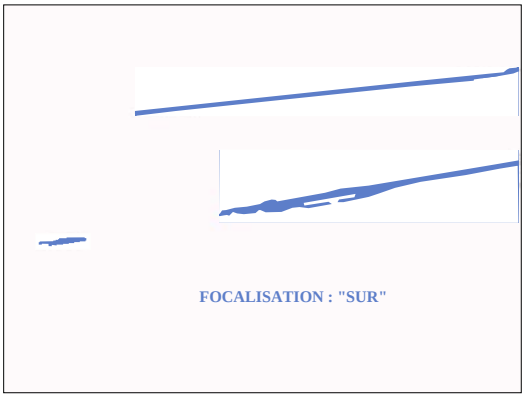


Fig. 5.6 – Image représentant les zones d'intérêt associées à l'objectif véhicule

créés. Chacun va demander la création d'un agent Localisation pour analyser la zone d'intérêt.

- sur les trois agents de type Localisation Véhicule, un seul trouve des infor-

mations significatives et représentatives du concept véhicule. Il extrait ces échantillons en question ;

– des agents Détection Véhicule exploitent ces échantillons en parallèle de façon à extraire le support. La probabilité des supports est modulée par les incertitudes déjà formulées sur les supports des routes. Rappelons que, le réseau bayésien qui permet de calculer la confiance dans un support est décrit sur la figure 5.7. Il intègre la variable Sup_p qui représente la confiance dans le

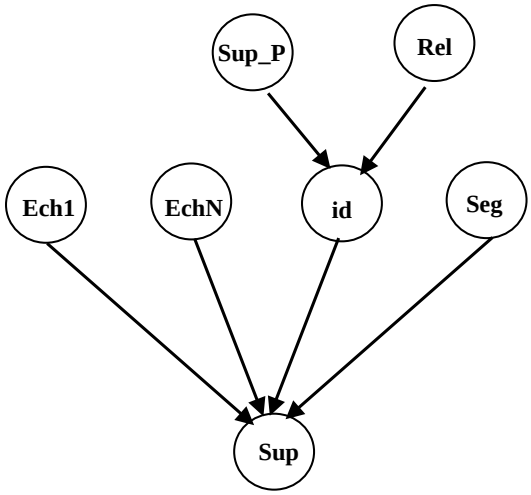


FIG. 5.7 – Réseau bayésien utilisé pour calculer la confiance dans un support.

support père, c'est-à-dire le support qui lui a donné naissance.

– Enfin, à partir du support cible, l'agent décision cherche à valider l'hypothèse avec une information mouvement qui n'est pas présente. Donc il prend la décision finale sans changer la probabilité obtenue.

Le résultat de segmentation obtenu est présenté sur la figure 5.8. Les supports route sont grisés alors que les supports véhicules apparaissent en vert. Cette couleur verte est représentative d'une probabilité d'appartenance au concept comprise entre 67 et 74%.

Le fait d'introduire des contraintes contextuelles élimine dans ce cas les fausses alarmes, sans diminuer les probabilités d'appartenance. Ceci s'explique du fait que la probabilité d'appartenance du support route est du même ordre de grandeur (exactement 70%) que les supports de type véhicule. Donc les probabilités restent stables.

Le graphe d'agents est présenté sur la figure 5.9. Il se décompose en trois branches principales à partir de l'agent Localisation Route (L_R). Deux de ces branches se terminent par un agent Localisation Véhicule (L_V). Aucune information significative relative au concept n'est trouvée sur le support Route, donc le traitement s'arrête de lui-même. Sur la troisième branche, trois échantillons véhicule sont localisés sur

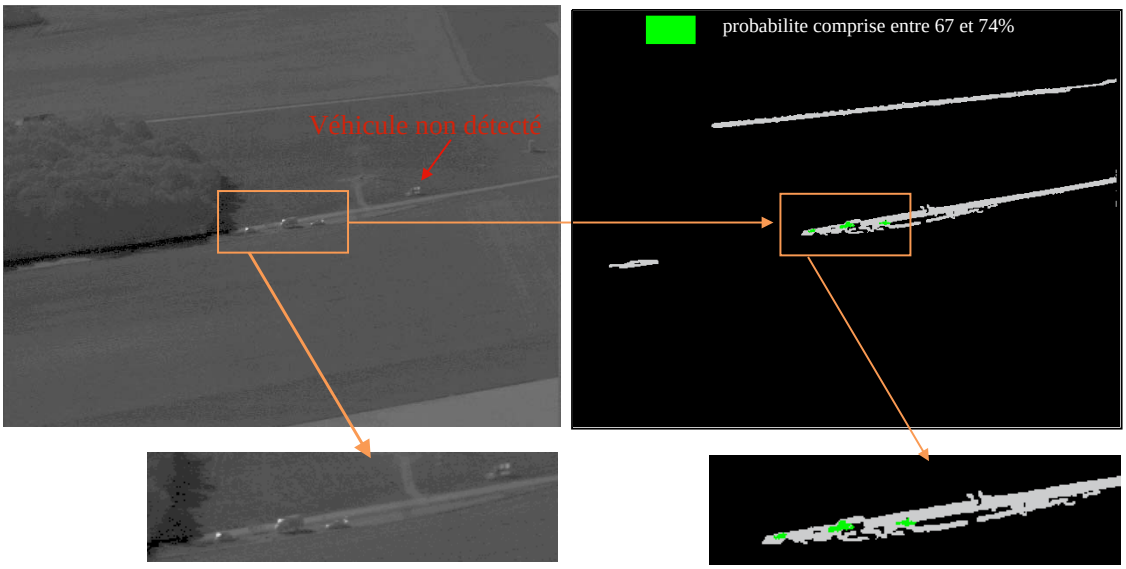


FIG. 5.8 – Segmentation des véhicules avec une règle d'exploration spatiale. À gauche on présente l'image originale. À droite l'image de segmentation contient trois supports de route en gris et trois supports de véhicule trouvés sur la route.

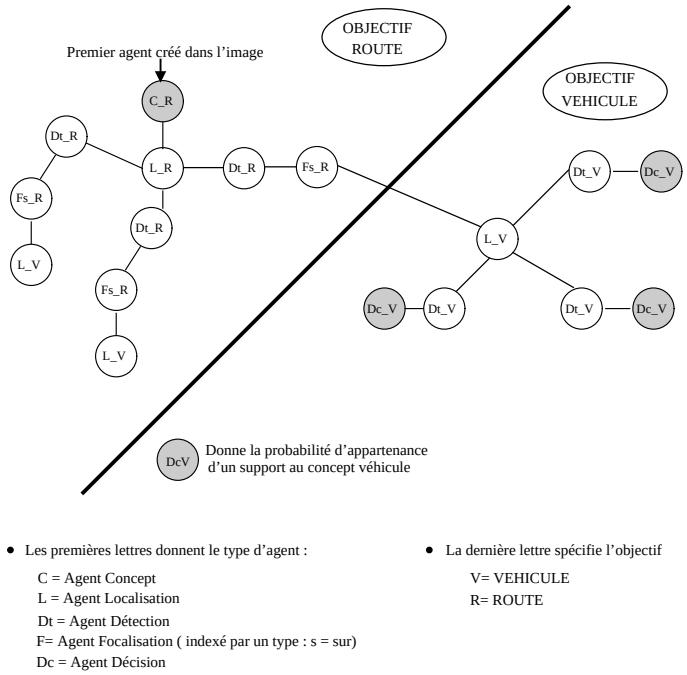


FIG. 5.9 – Graphe d'agents réalisant la détection de véhicules suivant la stratégie d'exploration " véhicules sur routes". Cinq types d'agent et deux concepts différents sont utilisés.

le support Route. On observe à nouveau une décomposition en trois branches.

Avec cette unique stratégie de résolution, un des quatre véhicules n'est pas détecté. Or, dans le cadre opérationnel, le problème est d'évaluer la menace pour prendre une décision. Aussi, le compromis entre le taux de fausses alarmes et le taux de non-détection n'est pas possible. Pour pallier cette difficulté, il faut diminuer la contrainte contextuelle en utilisant plusieurs règles de focalisation.

Plusieurs règles de focalisation

En augmentant le nombre de règles de focalisation, on augmente l'espace d'exploration où le système est susceptible de trouver l'information d'intérêt. Nous proposons un exemple avec deux règles de focalisation à savoir :

- chercher les véhicules sur les routes ;
- chercher les véhicules à côté des routes.

La base de connaissances utilisée est décrite page suivante et sur la figure 5.10. Dans cet exemple, le concept route comporte désormais deux stratégies de focalisation. Elle sont représentées en gras dans le fichier.

Par rapport à l'expérimentation précédente, on retrouve les mêmes types d'agents, de même et les mêmes buts. En revanche, la population et les régions d'exploration de l'image sont modifiées. À partir des supports de type route, différentes zones sont déterminées par les agents focalisation Route, pour chercher les véhicules. Celles-ci sont représentées sur la figure 5.11.

Les zones associées à la focalisation de type "sur" sont représentées en bleu, alors que les focalisations de type " à côté de" apparaissent en orange. En pratique, pour chaque support route, deux agents focalisation sont activés, chacun cherchant sa zone d'intérêt.

Pour réaliser la détection des véhicules dans cette image, les cinq types d'agents sont nécessaires à savoir, les agents Concept, Localisation, Détection, Focalisation et Décision. Ils se déclinent suivant deux objectifs possibles : Route et Véhicule. Ce sont les agents focalisation qui contrôlent à la fois l'exploration dans l'espace des buts et l'exploration dans l'espace image. Le résultat de segmentation obtenu est présenté sur la figure 5.12.

Quatre véhicules sont extraits. Trois se situent sur un des supports route, et le dernier est situé à côté de ce même support. Donc, par rapport à l'expérimentation précédente, on détecte tous les véhicules présents dans l'image. Le véhicule situé à côté de la route est caractérisé par une probabilité de l'ordre de 60%. Elle est nettement inférieure aux autres probabilités obtenues. Ceci est le résultat obtenu par le réseau bayésien qui considère les relations contextuelles. La table de probabilités conditionnelles entre le support Père et la relation de focalisation permet de renforcer ou non des hypothèses. C'est précisément le résultat que l'on observe. La

```
{DOMAINE}
RACINE :IMAGES/SOURCE/Vol/vol.%02d.gif
EXTENSION :gif :
DEBUT :0 :
NB :1 :
PAS :1 :
---------------------------------------
{BUT}
CONCEPT :4 :
---------------------------------------
{VALIDATION}
VAL :0.10
---------------------------------------
{CONCEPT 1}
NOM :VEGETATION :
TAILLE :1 :
L :1 :GRIS :2 :62 :81 :
L :1 :HOMOG :5 :0 :1 :
L :1 :SURFACE :0 :150 :10000 :

{CONCEPT 2}
NOM :CHAMP :
TAILLE :1 :
L :1 :GRIS :8 :72 :92 :
L :1 :HOMOG :8 :0 :10 :
L :1 :SURFACE :0 :1000 :30000 :
D :1 :DETECT :3 :0 :0

{CONCEPT 3}
NOM :VEHICULE :
TAILLE :1 :
L :1 :GRIS :1 :114 :255 :
L :1 :HOMOG :3 :1 :10 :
L :1 :SURFACE :0 :10 :100 :
D :1 :DETECT :1 :0 :0

{CONCEPT 4}
NOM :ROUTE :
TAILLE :1 :
L :1 :HOUGH :1 :0.1 :0.5 :
L :1 :GRIS :4 :80 :255 :
L :1 :SURFACE :1 :250 :4000 :
D :1 :DETECT :2 :0 :0
F :1 :FOCUS :S :3
F :1 : FOCUS :A :3
```

FIG. 5.10 – Exemple de fichier de configuration

probabilité d'observer un véhicule sur la route est élevée, alors qu'elle est plus faible de localiser un véhicule près de la route.

Le graphe d'exploration associé à cette segmentation est présenté sur la figure 5.13. Par rapport à l'expérimentation précédente, il y a un dédoublement des branches de focalisation. Ceci signifie que deux agents focalisation sont activés lorsqu'un support route est trouvé. Il y a une focalisation sur le support route, comme dans l'expérimentation précédente, qui est complétée par focalisation à côté des supports route.

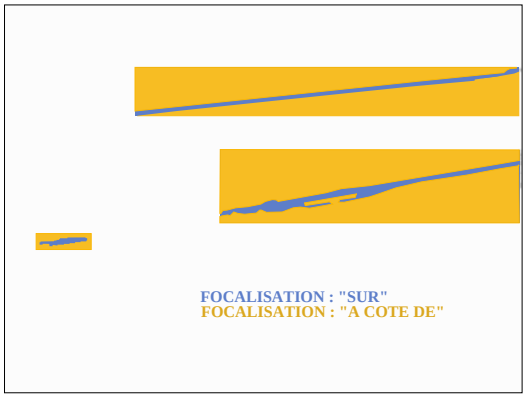


FIG. 5.11 – Ensemble des zones d'intérêt définies par les agents Focalisation. En orange, sont représentées les zones issues d'une focalisation de type "sur", et en bleu celles relatives à une focalisation du type "à côté de".

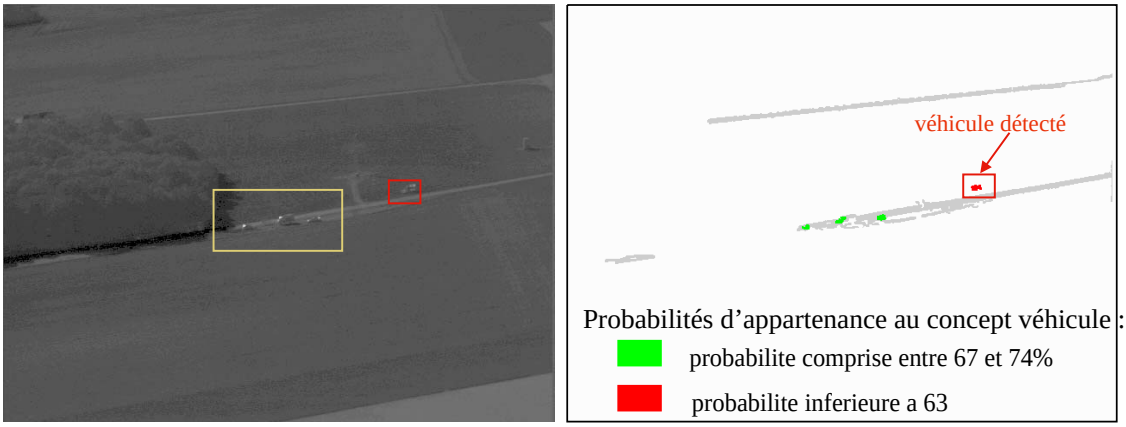


FIG. 5.12 – Segmentation des véhicules avec deux stratégies de focalisation. À gauche, on présente l'image originale et à droite le résultat de la segmentation. Quatre véhicules sont détectés : trois sur la route et un à côté de la route.

Nous nous sommes intéressés jusqu'à présent à l'intérêt de l'exploration spatiale de l'image. Suivant les choix effectués, les contraintes et les probabilités évoluent. En effet, l'espace d'exploration et les probabilités d'appartenance varient.

L'ensemble des résultats présentés s'intéresse aux traitements intra-image, nous allons maintenant considérer une séquence d'images et présenter notamment les principes d'incrémentalité et d'adaptativité.

5.5 Exploitation d'une séquence d'images

Autant dans le paragraphe précédent nous avons choisi d'insister sur la population d'agents, son organisation et sa distribution, autant dans ce paragraphe, nous

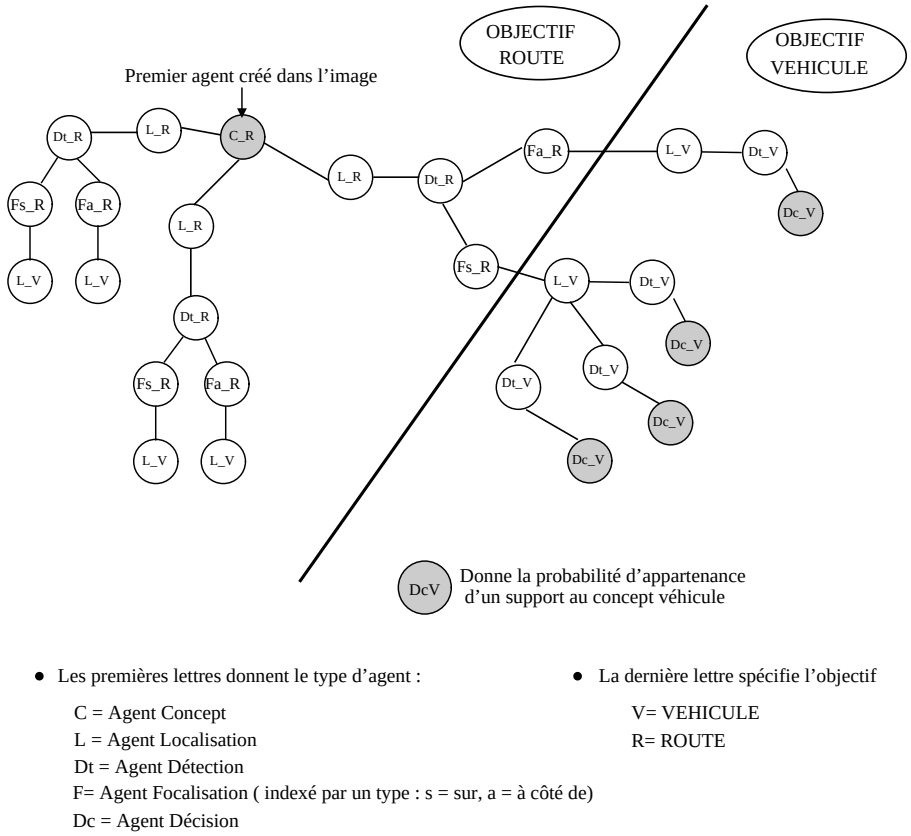


FIG. 5.13 – Graphe d'agents relatif à la détection de véhicules lorsque deux règles de focalisation sont spécifiées. L'objectif initial est route. Le système recherche les véhicules sur les routes et à côté des routes.

allons montrer les interactions des agents avec les autres modules du système, tels que le modèle du monde et la base de connaissances. Nous détaillerons principalement le principe d'adaptativité et la boucle fusion d'informations.

5.5.1 Les capacités d'adaptation du système

Comme nous l'avons décrit au cours du précédent chapitre notre système est doté de différentes capacités d'adaptation. On distingue principalement :

– l'adaptation locale, qui est notamment présente au sein de l'agent Détection puisqu'il utilise les statistiques associées à l'échantillon courant pour effectuer sa segmentation.

– l'adaptation globale est présente sous deux formes différentes :
  • adaptation de la population d'agents : par exemple, la présence de mouvement dans l'image peut contraindre le système à demander une analyse complémentaire dans les régions concernées. Ainsi de nouveaux agents sont

créés pour répondre à cette demande.

• adaptation des structures de données : sur une séquence d'images, le système est capable de gérer des perturbations grace à une mise à jour de la base de connaissances et une recopie des échantillons. Les éventuelles perturbations provoquent toutefois une évolution du modèle du monde avec une diminution des probabilités. Nous montrerons également l'influence du seuil de validation.

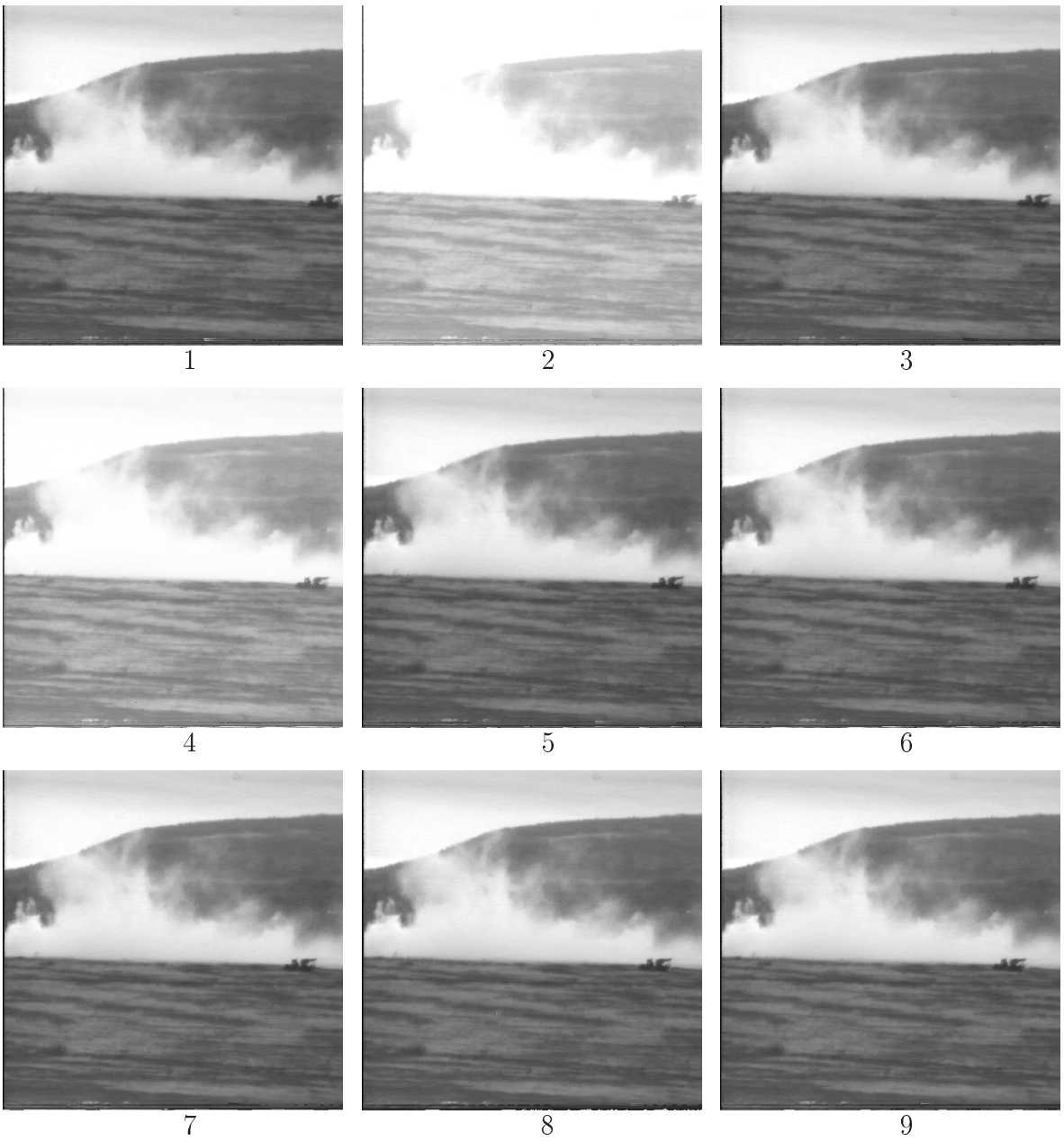


Fig. 5.14 – Séquence de référence

Nous donnerons dans ce paragraphe une illustration des différentes formes d'adaptation globale du système. Nous avons choisi de présenter les résultats sur la séquence d'images numérotée 3 (voir paragraphe 5.3). Cette séquence est composée de 9 images. Nous avons introduit une perturbation (augmentation uniforme des niveaux de gris) dans les images numérotées un et trois de façon à mieux mesurer le caractère adaptatif du système. Une vue de la séquence ainsi constituée est présentée sur la figure 5.14

5.5.2 L'adaptation de la population d'agents

Comme nous l'avons précisé dans le paragraphe 4.6.5, l'agent mouvement réalise la fonction de veille dans le système. Tout objet en mouvement est repéré par l'agent mouvement qui demande ensuite un complément d'informations avec la création d'agents spécialisés.

Nous allons détailler complètement un exemple, afin de bien comprendre le principe de fonctionnement du système. Suivant la même philosophie que précédemment, nous avons généré avec notre interface graphique et à partir de la première image de la séquence, notre base de connaissances. Celle-ci est présentée sur la figure 5.15.

Le concept véhicule est décrit dans la base de connaissances avec un champ mouvement. Aussi, à chaque nouvelle image, l'agent Concept demande l'activation d'un agent mouvement en parallèle de l'exploration spatiale de l'image. Cette exploration spatiale se fait suivant une stratégie de type "sur". Le système va rechercher les véhicules sur les champs.

Le résultat de la segmentation au sens du mouvement, obtenu pour la première image de la séquence, est donné sur la figure 5.16. Le véhicule est étiqueté, mais il y a quatre autres régions représentatives du concept fumée qui sont extraites. Chaque région apparaît comme une cible potentielle, à partir de laquelle on va poursuivre l'analyse. Ce cheminement est représenté dans le graphe d'agent de la figure 5.17. Ce graphe nécessite quelques explications car il diffère des graphes précédents. En effet, on distingue clairement deux branches :

- la première est relative à l'exploration spatiale de l'image. Elle se situe à gauche du graphe. Elle commence par la localisation et la détection des champs dans l'image. Pour chaque support trouvé, une règle de focalisation est appliquée : rechercher les véhicules sur les supports champs. Les régions déjà étiquetées par la branche mouvement avec le label véhicule ne sont donc pas explorées ;
- la deuxième est relative à la recherche du mouvement. Elle se situe à droite sur le graphe. Cette branche est ensuite éclatée en cinq sous-branches puisque le résultat de la segmentation au sens du mouvement a permis d'extraire cinq régions. Mais un seul de ces supports est identifié finalement comme une cible par un agent de Décision.

```
{DOMAINE}
RACINE :IMAGES/SOURCE/Mif1/mif.%02d.gif
EXTENSION :gif :
DEBUT :0 :
NB :10 :
PAS :1 :
---------------------------------------
{BUT}
CONCEPT :1 :
---------------------------------------
{VALIDATION}
VAL :0.10
---------------------------------------
{CONCEPT1}
NOM :CHAMP :
TAILLE :1 :
L :0 :HOUGH :0 :0.1 :0.5 :
L :1 :GRIS :5 :40 :120 :
L :1 :HOMOG :9 :0 :3 :
L :1 :SURFACE :0 :250 :3000000 :
D :1 :DETECT :5 :0 :0
F :1 :FOCUS :S :3

{CONCEPT2} NOM :FUMEE :
TAILLE :1 :
L :0 :HOUGH :0 :0.1 :0.5 :
L :1 :GRIS :2 :140 :255 :
L :1 :HOMOG :6 :0 :1 :
L :1 :SURFACE :0 :100 :100000 :

{CONCEPT3}
NOM :VEHICULE :
TAILLE :2 :
L :1 :GRIS :5 :10 :80 :
L :1 :HOMOG :2 :3 :15 :
L :1 :SURFACE :0 :150 :1000 :
M :1 :MOUV :3 :255 :
```

FIG. 5.15 – Fichier de configuration relatif à la première image de la séquence de référence.

FIG. 5.16 – Segmentation des zones en mouvement : à gauche est présentée l'image de référence et à droite, le résultat issu de la segmentation.

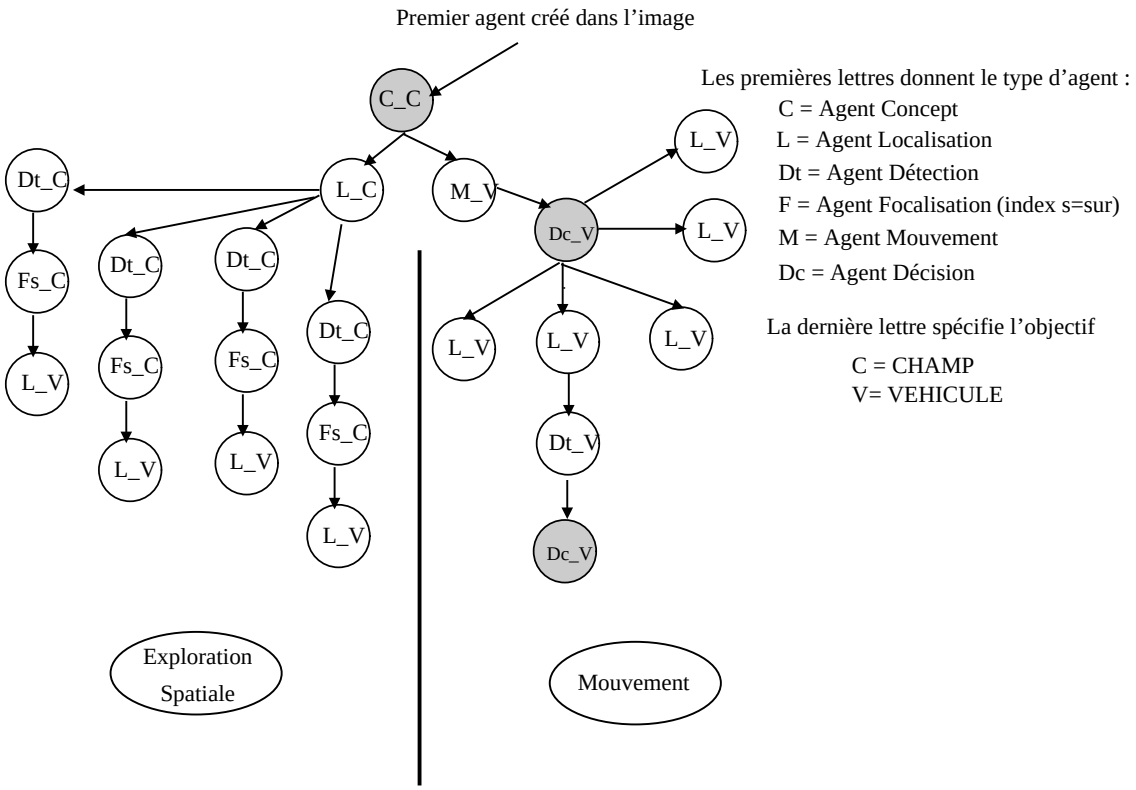


FIG. 5.17 – Graphe d'agents en présence du mouvement. Il représente la population d'agents mise en œuvre pour trouver les véhicules sachant, que l'exploration spatiale est accompagnée de la recherche des informations mouvement.

Dans cet exemple, l'objet est détecté par la boucle liée à la recherche du mouvement. Ceci est dû à la rapidité d'exécution de cette boucle par rapport à l'exploration spatiale. Mais, dans l'absolu, il se peut que le véhicule en mouvement soit trouvé par l'une ou l'autre branche. Les temps d'execution de chacun des agents est présenté sur la figure 5.18.

Dans le graphe d'agent, on constate que deux agents Décision sont activés au cours du traitement d'une image. Le premier est activé sur demande de l'agent mouvement, il regarde si la zone en mouvement a déjà été étiquetée avec le label véhicule. Dans l'affirmatif, cette branche se serait interrompue et aurait fourni la probabilité d'appartenance au concept véhicule. Dans le cas qui nous intéresse, il s'avère que les régions en mouvement ne sont pas encore étiquetées par le concept véhicule. Aussi, une analyse plus approfondie est demandée, ce qui se traduit par la création de plusieurs agents Localisation.

La prise en compte de l'information mouvement permet d'augmenter la probabilité d'appartenance du support au concept véhicule. La probabilité obtenue est de 79%, alors que sans cette information mouvement la probabilité est de 65%.

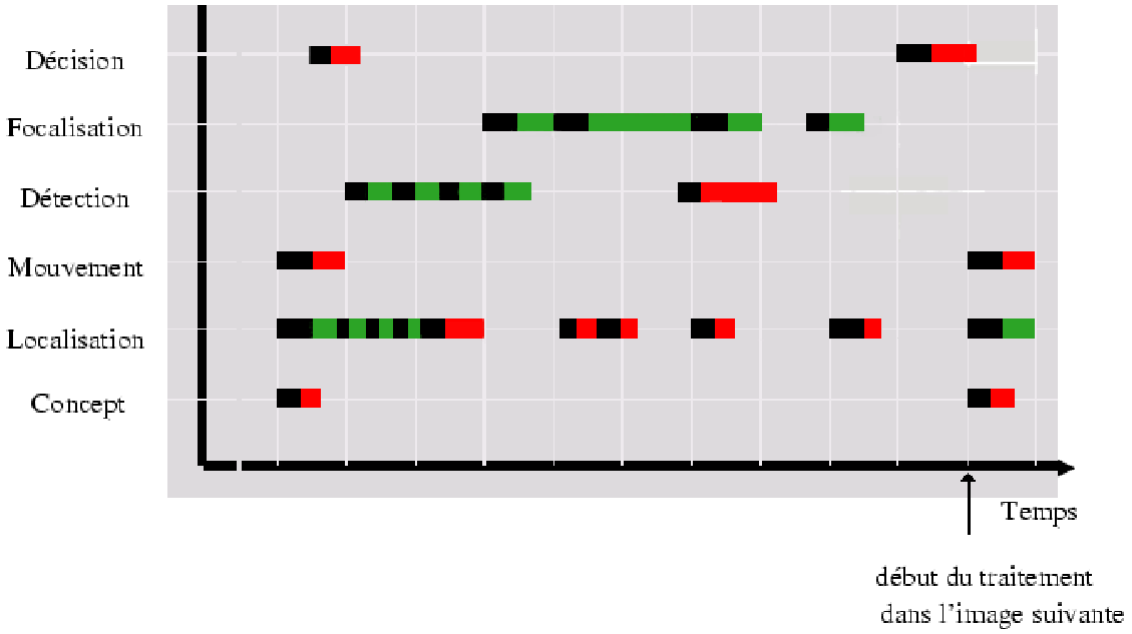


FIG. 5.18 – Répartition des temps de calculs entre agents : en vert sont représentés les agents liés à l'exploration de l'image et en rouge les agents associés à la boucle mouvement.

L'agent Mouvement est un agent de veille qui permet de prendre plus rapidement une décision sur l'éventualité d'une menace. En revanche cette analyse n'altère en rien l'intérêt de l'exploration spatiale qui segmente alors les cibles non détectées par la branche mouvement soit pour des raisons de temps de calcul (suivant par exemple la charge de la machine...), soit parce que le véhicule en question est à l'arrêt.

Suivant l'information disponible et la rapidité d'exécution, les chemins pour segmenter la cible en question peuvent être modifiés. Le système peut utiliser un fonctionnement exploratoire basé sur des relations contextuelles ou sur une information forte telle que le mouvement. Le système s'adapte à la situation rencontrée.

5.5.3 Adaptation de la base de connaissance et du modèle du monde

Comme nous l'avons décrit dans le paragraphe 4.4.3, notre système se caractérise également par une accumulation des connaissances au cours du temps. Cette accumulation des connaissances entre les instants t et t+1 s'effectue en deux étapes. Dans un premier temps il y a une recopie des échantillons associés à un concept fixe et possédant une confiance suffisante (supérieure à un seuil de validation). Et dans un deuxième temps, il y a une mise à jour de la base de connaissances à partir des statistiques associées à ces échantillons. Cette phase de mise à jour est prise en

```
{DOMAINE}
RACINE :IMAGES/SOURCE/Mif1/mif.%02d.gif
EXTENSION :gif :
DEBUT :0 :
NB :10 :
PAS :1 :
---------------------------------------
{BUT}
CONCEPT :1 :
---------------------------------------
{VALIDATION}
VAL :0
---------------------------------------
{CONCEPT1}
NOM :CHAMP :
TAILLE :1 :
L :0 :HOUGH :0 :0.1 :0.5 :
L :1 :GRIS :5 :40 :120 :
L :1 :HOMOG :9 :0 :3 :
L :1 :SURFACE :0 :250 :3000000 :
D :1 :DETECT :5 :0 :0

{CONCEPT2} NOM :FUMEE :
TAILLE :1 :
L :0 :HOUGH :0 :0.1 :0.5 :
L :1 :GRIS :2 :140 :255 :
L :1 :HOMOG :6 :0 :1 :
L :1 :SURFACE :0 :100 :100000 :

{CONCEPT3}
NOM :VEHICULE :
TAILLE :2 :
L :1 :GRIS :5 :10 :80 :
L :1 :HOMOG :2 :3 :15 :
L :1 :SURFACE :0 :150 :1000 :
```

FIG. 5.19 – Fichier de configuration relatif à la première image de la séquence.

charge à chaque nouvelle image par l'agent Concept.

5.5.4 Influence d'une perturbation

Dans ce paragraphe nous proposons une illustration de ces aspects de mise à jour de la connaissance. Pour cela nous avons choisi de travailler avec la séquence présentée sur la figure 5.14, et d'évaluer à partir de celle-ci la capacité d'adaptation du système face à la perturbation introduite.

À partir de la première image de cette séquence, le fichier de configuration a été généré. Celui-ci est présenté sur la figure 5.19.

L'objectif est de segmenter les supports associés au concept champ. Pour cela, les attributs de type niveaux de gris, homogénéité et surface, sont utilisés. L'opérateur humain spécifie que tous les échantillons doivent être conservés pour la mise à jour des connaissances par conséquent, le seuil de validation est fixé à 0. Le résultat de

segmentation obtenu sur les trois premières images de la séquence est présenté sur la figure 5.20.

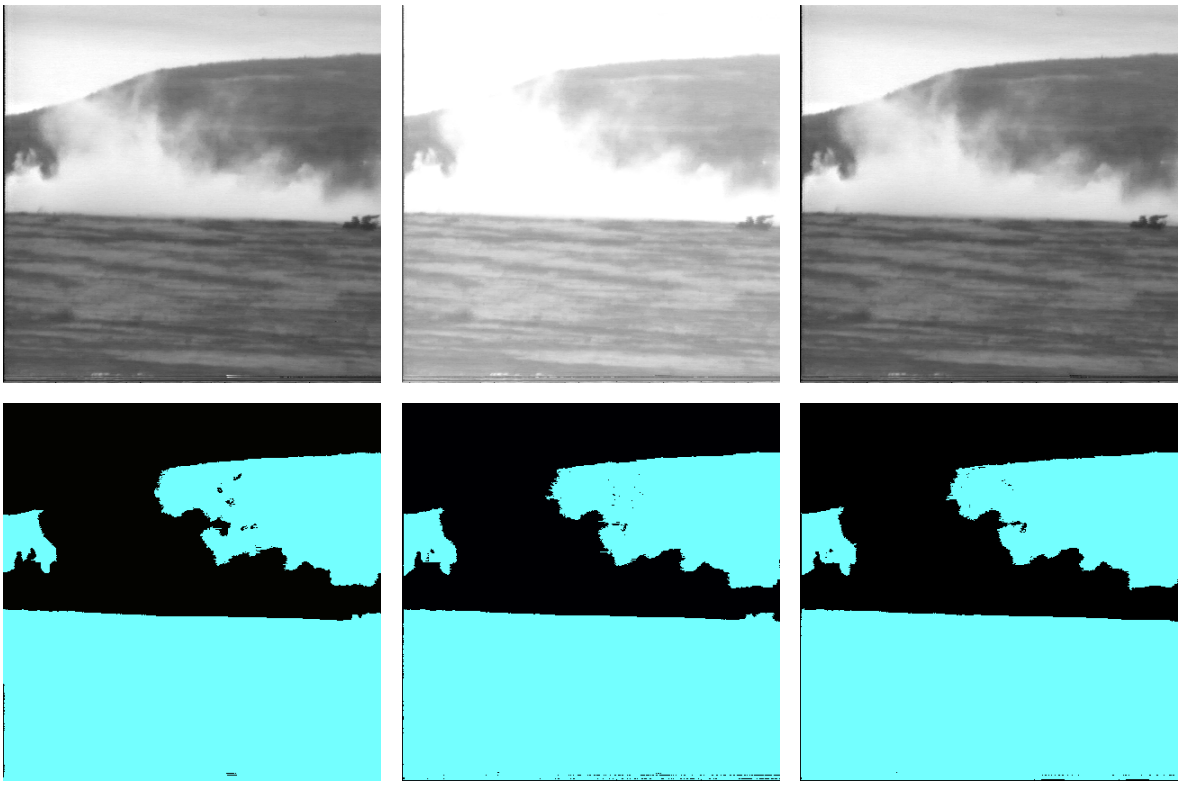


FIG. 5.20 – Résultat de segmentation

Il s'avère que le système effectue une segmentation correcte malgré la forte perturbation introduite dans la deuxième image. Les trois supports champs sont extraits.

Ce résultat est dû au caractère adaptatif de l'algorithme. L'explication en est très simple. En effet, pour analyser l'image perturbée, l'agent Concept commence par recopier les échantillons de l'image précédente qui ont une confiance suffisante, et fait une mise à jour de la base de connaissances. Ensuite il demande la création d'un agent localisation qui réalise une analyse statistique pour mettre à jour les données attachées aux échantillons. Ces nouvelles informations sont ensuite utilisées par l'agent Détection pour faire sa croissance de régions. C'est précisément l'analyse statistique réalisée par l'agent Localisation qui permet une adaptation par rapport au contexte, et une robustesse vis-à-vis des perturbations éventuelles.

Maintenant que nous avons observé le résultat de segmentation, il est intéressant de continuer cette analyse de la séquence d'images, en regardant les grandeurs statistiques associées à chaque support, ainsi que les variations des incertitudes.

Afin de faciliter la lisibilité des résultats nous avons choisi de numéroter les supports de la façon décrite sur la figure 5.21.

FIG. 5.21 – Numérotation des supports d'objets

Les résultats obtenus pour l'attribut niveau de gris sont présentés dans le tableau 5.14. Ils sont indexés par le numéro du support et le numéro de l'image. De plus, pour chaque support, dans chaque image, six paramètres sont spécifiés :

- Nb Pixel = le nombre de pixels constituant le support ;
- Nb Ech = le nombre d'échantillons ayant participé à la segmentation du support ;
- Moyenne = moyenne des niveaux de gris des pixels appartenant au support ;
- Valeur Max = niveau de gris maximum trouvé dans le support ;
- Valeur Min = niveau de gris minimum trouvé dans le support ;
- Estimation = coefficient d'appartenance du support au concept Champ.

À partir de ce tableau de données, une analyse s'impose :

Les informations statistiques

D'un support à l'autre et d'une image à l'autre, les grandeurs statistiques obtenues, à savoir la moyenne et les extrema des niveaux de gris (Valeur Max et Valeur Min) sont très différentes. Par exemple pour le support $n^0 3$, la moyenne varie de 105 dans l'image 1, à 180 dans l'image 2. Elle met en valeur la perturbation introduite. Cette grande variabilité se retrouve d'ailleurs entre les images $n^0 2$ et $n^0 3$.

L'ensemble des statistiques obtenues pour chaque image permet de mettre à jour la base de connaissances et plus spécifiquement l'intervalle de validation de l'attribut. En l'occurrence, avec les extrema Valeur Max et Valeur Min le système pourra modifier l'intervalle de validation de l'attribut niveaux de gris. Cela se fait par un moyennage, pondéré par le nombre de pixels, et uniquement pour les supports qui ont une confiance supérieure au critère de validation.

Le nombre d'échantillons

Le nombre d'échantillon donne une information sur le processus de construction du support obtenu. Il précise notamment si le support a été créé par simple croissance de régions, ou si un procédé de fusion a été ajouté.

		Support $n^0 1$	Support $n^0 2$	Support $n^0 3$
Image $n^0 0$	– Nb pixel – Nb Ech – Moyenne – Valeur Max – Valeur Min – Estimation	– 2667 – 1 – 103.138 – 71 – 123 – 61.3804	– 23800 – 2 – 97.46 – 49 – 126 – 52.8577	– 76335 – 1 – 105.64 – 78 – 125 – 71.3112
Image $n^0 1$	– Nb pixel – Nb Ech – Moyenne – Valeur Max – Valeur Min – Estimation	– 2933 – 1 – 186.561 – 149 – 217 – 20.8702	– 25800 – 2 – 189 – 160 – 213 – 42.92	– 76681 – 1 – 180.015 – 109 – 238 – 39.1423
Image $n^0 2$	– Nb pixel – Nb Ech – Moyenne – Valeur Max – Valeur Min – Estimation	– 2706 – 1 – 103.858 – 68 – 141 – 52.5831	– 24542 – 2 – 106.5339 – 78 – 127 – 69.9107	– 75871 – 1 – 97.7013 – 19 – 177 – 59.9037
Image $n^0 3$	– Nb pixel – Nb Ech – Moyenne – Valeur Max – Valeur Min – Estimation	– 3120 – 1 – 161.424 – 114 – 219 – 21.6367	– 24378 – 2 – 163.46 – 130 – 195 – 12.13577	– 76704 – 1 – 152.389 – 82 – 244 – 19.9492
Image $n^0 4$	– Nb pixel – Nb Ech – Moyenne – Valeur Max – Valeur Min – Estimation	– 3126 – 1 – 104.936 – 69 – 173 – 34.5756	– 24821 – 2 – 107 – 79 – 135 – 52.5136	– 76506 – 1 – 97.907 – 49 – 200 – 74.331
Image $n^0 5$	– Nb pixel – Nb Ech – Moyenne – Valeur Max – Valeur Min – Estimation	– 3341 – 1 – 104.95 – 63 – 196 – 35.4027	– 26370 – 2 – 106.965 – 79 – 136 – 65.7741	– 76480 – 1 – 98.0214 – 44 – 204 – 73.2798
Image $n^0 6$	– Nb pixel – Nb Ech – Moyenne – Valeur Max – Valeur Min – Estimation	– 3735 – 1 – 105.425 – 66 – 219 – 48.6496	– 26481 – 2 – 107.306 – 77 – 141 – 65.5402	– 76668 – 1 – 97.8532 – 24 – 203 – 73.6107

FIG. 5.22 – Description des différents supports présents dans la séquence d'images.

L'incertitude

La probabilité d'appartenance d'un support à un concept varie au cours de l'analyse de la séquence d'images. On constate notamment deux brusques variations pour les images 1 et 3 qui correspondent aux images perturbées. Il existe à ce moment là, un fort contraste entre les informations stockées dans la base de connaissances et la mesure extraite de l'image. Ce qui explique les faibles probabilités d'appartenance. À la suite de ces deux perturbations, les probabilités des supports augmentent progressivement pour finalement se stabiliser. Les échantillons deux et trois obtiennent respectivement une probabilité d'appartenance au concept champ de 65 et 73 %. La probabilité associée à l'échantillon 1 reste peu stable. Ceci s'explique notamment par le faible nombre de pixels considérés.

Cette expérience met en valeur le caractère adaptatif du système. En fonction des statistiques locales, l'agent Détection adapte ses paramètres. Cette adaptation locale associée à la mise à jour de la base de connaissances au cours du temps, permet d'être robuste à de fortes perturbations. L'introduction de ces perturbations garde toutefois un impact très important sur les valeurs de probabilités puisque celles-ci traduisent l'adéquation entre la base de connaissances et les mesures issues de l'image.

5.5.5 Influence du seuil de validation

Au cours de ces différentes expérimentations, nous avons remarqué que les graphes d'agents variaient suivant les spécifications de la base de connaissances (règles de focalisation), mais également suivant les informations disponibles et valides. Notre système met à jour cette base, uniquement si les probabilités attachées aux supports sont suffisantes. Et c'est précisément ce lien entre stratégies et probabilités qui nous intéresse maintenant. La question est donc d'évaluer l'impact des probabilités sur la population d'agents.

À plusieurs reprises nous avons parlé du critère de validation. Ce paramètre spécifié dans la base de connaissances définit le seuil d'acceptation à la fois pour la recopie des échantillons et la mise à jour de la base de connaissances. Il précise la limite entre l'information obtenue qui sera conservée ou rejetée. Aussi, nous allons évaluer son influence sur le système.

Nous illustrons ce résultat sur un exemple très simple, où le critère de validation est modifié. Reprenons le tableau 5.22, nous constatons que les probabilités associées aux échantillons sont respectivement de 61, 52 et 71, pour les supports 1, 2 et 3, dans l'image 0. Lorsque le seuil de validation est fixé à 0, comme précédemment, tous les échantillons qui ont donné naissance à ces supports sont recopiés dans l'image suivante, et sont utilisés pour faire une mise à jour de la base de connaissances. Maintenant lorsque le seuil de validation est augmenté, le système ne sélectionne que les meilleurs résultats pour effectuer la mise à jour de la base de connaissances et du modèle du monde. Les résultats obtenus pour la segmentation de l'image n^0 1

sont présentés sur la figure 5.23

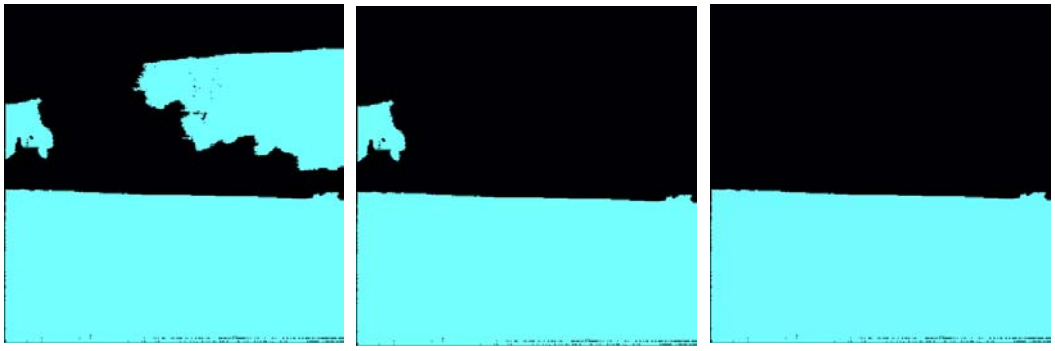


FIG. 5.23 – Segmentation de la première image avec différents seuils de validation. À gauche, le seuil est fixé à 0, au centre, il est de 55 et enfin à droite, il est de 65.

Rappelons quand même, que l'image 1 comporte une perturbation. Donc, si le procédé de recopie des échantillons n'est pas utilisé, le système n'arrive pas à retrouver la région représentative du concept avec les seules informations de la base de connaissances. En effet, on constate que plus le seuil de validation est élevé, moins le nombre d'échantillons recopié est important. Ceci se traduit par une segmentation partielle de l'image. Ceci signifie plusieurs choses :

- lors d'une forte perturbation, comme sur l'image $n^0 1$, ce ne sont pas les données spécifiées dans la base de connaissances, qui permettent la segmentation, mais davantage la prise en compte de l'accumulation des connaissances. En effet, si la recopie des échantillons n'est pas effectuée avec une analyse statistique locale, le système n'arrive pas suivant le principe de localisation à extraire une région de l'image.

- plus le seuil de validation est élevé, plus les contraintes sur les données sont fortes. En effet, il faut une confiance très importante pour propager l'information. Pour un seuil de validation à 50, un seul échantillon est utilisé pour la mise à jour de la base de connaissances et la recopie des données. Ainsi, suivant les informations extraites, à l'instant t-1, les résultats à l'instant t sont différents. Et donc la population d'agents s'en trouve modifiée, puisqu'elle s'organise en fonction de l'information disponible.

Il existe ainsi une dépendance mutuelle entre les stratégies et les probabilités.

5.6 Conclusion

Plusieurs caractéristiques du système sont mises en évidence au travers de ces expérimentations.

La distribution du système se décline sous plusieurs formes :
- distribution des traitements qui décompose la tâche de vision en sous-tâches (localisation, détection, focalisation, mouvement et décision) ;
- distribution spatiale : les agents sont situés dans l'image ;
- distribution suivant l'objectif courant.

Cette distribution permet d'une part une grande modularité et d'autre part une capacité à s'adapter au contexte. Ceci est représenté au travers des différents graphes d'agents. Suivant les informations trouvées, la population d'agents évolue au cours du temps et construit de façon incrémentale le modèle du monde.

Le fait d'ajouter une connaissance sous forme d'un modèle contextuel (stratégie de focalisation) diminue considérablement le nombre de fausses alarmes. En revanche cette connaissance doit être spécifiée par l'opérateur *a priori*. L'agent Mouvement quant à lui est doté d'un comportement de veille, qui lui confère le pouvoir de détecter très vite les zones d'intérêt, dans le cas de véhicules en mouvement. Il y a une véritable complémentarité entre le mouvement qui va rechercher les véhicules en mouvement en demandant une analyse très rapidement, et l'exploration spatiale qui construit progressivement son analyse de façon à compléter la détection.

L'incertitude est représentée sous la forme d'une probabilité d'appartenance à un concept. Elle varie suivant les stratégies utilisées (la filiation) et la mesure obtenue. Inversement, les stratégies et même la mise à jour des connaissances s'adaptent suivant les valeurs de probabilités. Cette mise à jour est paramétrée dans le système notamment avec le seuil de validation.

L'adaptation du système est réalisée à plusieurs niveaux. Il y a une adaptation :

- des paramètres des agents. Par exemple l'agent Détection utilise les statistiques locales pour effectuer sa croissance de région.
- de l'organisation du système selon les connaissances spécifiées dans la base de connaissances et les informations collectées dans l'image (par exemple le mouvement).
- du modèle du monde suivant les confiances obtenues
- de la base de connaissances avec la mise à jour des intervalles de validation suivant les mesures obtenues.

Ces qualités d'adaptation sont importantes, et surtout dans notre approche. En effet, initialement le système est très lié au fichier de configuration. Et on est précisément conscient de ses limites vis-à-vis d'un apprentissage complet. Aussi notre objectif a été de favoriser l'accumulation des connaissances au cours du temps, plutôt que de se baser sur un *a priori* approximatif.

Lorsque le traitement est terminé, le modèle du monde contient notamment les zones où les véhicules sont présents, et des informations statistiques associées à chaque support. L'étape suivante est maintenant de donner un nom à ce véhicule en

question. C'est l'objet du chapitre suivant concernant les reconnaissance de cibles.

Chapitre 6

Vers la reconnaissance des cibles

Sommaire

6.1 Introduction **145**
6.2 Présentation du système de reconnaissance **145**
6.3 Description des agents **147**
6.4 Les attributs pour la reconnaissance **148**
6.5 Apprentissage **152**
6.6 La fonction utilité **158**
6.7 Conclusion **160**

Résumé

Ce chapitre propose une approche pour la phase de reconnaissance basée sur la sélection d'attributs. L'objectif est d'utiliser les informations pertinentes pour réaliser une discrimination efficace. La sélection de variables apparaît comme une solution pour réduire la quantité d'informations à traiter et améliorer les performances. Nous proposons une approche basée sur la prédiction et la vérification d'hypothèses pour tirer profit des informations de position disponibles et faciliter cette reconnaissance.

Nous avons gardé la même philosophie par rapport à la phase de détection. On s'intéressera donc au principe d'incrémentalité, d'adaptativité, et de distribution. En revanche, nous avons intégré la notion d'utilité dans le choix des attributs. Ceci va se traduire par une légère modification de la représentation de l'incertitude. Il s'avère qu'une version spécifique des réseaux bayésiens, appelée diagramme d'influences, permet une adaptation à ce type de problème.

Nous illustrerons nos propos, en donnant quelques résultats obtenus à partir d'une base d'apprentissage, constituée d'images de synthèse.

6.1 Introduction

Le protocole utilisé en détection répond, jusqu'à présent, aux principes de localisation d'échantillons, de focalisations successives, de détection, de fusion d'informations, d'adaptativité et d'incrémentalité. Des régions de l'image sont progressivement segmentées et donnent une hypothèse sur la présence de cibles et d'éléments de contexte. L'objectif de cette phase de détection est d'extraire le support des cibles éventuellement présentes dans la séquence d'images et de calculer un ensemble d'informations statistiques, contextuelles, physiques... L'objectif désormais est de poursuivre cette analyse et de déterminer l'identité du support en question.

L'information extraite va désormais être exploitée en vue de spécifier le type d'un élément (le nom de la cible), non plus suivant une philosophie de caractérisation, mais davantage suivant un procédé de discrimination. Nous présenterons dans ce chapitre le principe de fonctionnement aux travers de nouveaux agents. Nous discuterons le choix des modèles et leur intégration dans le système.

La reconnaissance peut être perçue comme un problème de sélection d'attributs. Toutes les informations accumulées ne sont pas forcément pertinentes, elles peuvent correspondre à du bruit, être non informatives, corrélées ou non pertinentes pour notre problématique. La sélection de variables réduit la quantité d'informations à traiter et améliore les performances avec la disparition des variables non pertinentes.

Il existe une très grande variabilité intra et inter-classe, comme nous l'avons montré dans l'introduction. La probabilité d'erreur est potentiellement importante, d'où la nécessité de contraindre l'analyse par des informations supplémentaires (contexte, position) et des choix d'attributs pertinents. Le procédé utilisé est basé sur une approche de type prédiction-vérification, à laquelle on a associé la notion d'utilité d'attribut.

6.2 Présentation du système de reconnaissance

La phase bas niveau de l'architecture fournit le support de la cible, une hypothèse et une liste d'informations associées (moyenne, variance, relation contextuelles...). Ces données peuvent contribuer à la reconnaissance de la cible, sans pour autant être suffisantes. Par exemple :

- l'environnement informe sur la nature de la cible. On a plus de chance de trouver un véhicule sur la route qu'un navire sur la route. De même, il est plus probable d'observer un avion dans le ciel qu'un char dans le ciel. Aussi, suivant le contexte d'apparition des hypothèses peuvent être renforcées.

- le mouvement peut également nous donner une information précieuse. Un avion se déplace beaucoup plus vite qu'un char.

Toutefois, même si ces exemples paraissent tout à fait pertinents, la réciproque n'est pas forcément vérifiée. C'est-à-dire qu'un faible mouvement dans une image, ne signifie pas l'absence d'avions. Il est tout à fait plausible que le char arrive droit sur le capteur, ce qui se traduit par un mouvement apparent faible. Par ces quelques exemples, on souligne l'intérêt de la discrimination par rapport à la caractérisation. L'objectif est de trouver le ou les paramètres qui permettront de prendre une décision satisfaisante, dans un contexte donné.

Il apparaît clairement que, suivant la nature de la cible, son contexte d'apparition ou encore les conditions d'acquisition, la perception est différente. Ainsi, pour la phase de reconnaissance l'ensemble des attributs pertinents pour la caractérisation peut s'en trouver modifié. Suivant la position, la cible présente notamment des caractéristiques de symétrie et de compacité différentes. L'exemple présenté sur la figure[1] 6.1 montre la diversité des apparences suivant que l'on observe un mirage 2000 de face, de dessus ou de profil. Trouver un vecteur d'attributs caractéristique qui puisse reconnaître cette cible quelque soit la position ne semble pas trivial.

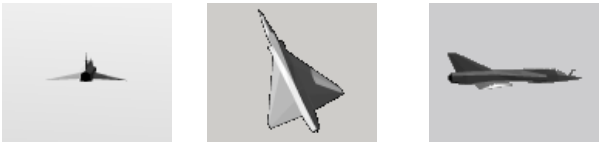


Fig. 6.1 – Différentes vues d'un mirage 2000. À gauche, le mirage est vu de dos, au centre il est observé de dessus et enfin à droite, il est positionné de profil.

Face à cette difficulté nous avons voulu contraindre ce problème en utilisant toutes les informations à notre disposition :

- l'environnement peut apporter des informations sur la nature de la cible, et parfois même sur son attitude par rapport au capteur
- le type de mission permet de formuler des hypothèses sur les positions des objets dans leur contexte. Par exemple dans une mission air-sol, il est probable d'observer des véhicules de dessus, plutôt que de profil.

L'objectif n'est donc plus d'utiliser un vecteur d'attributs par type de cibles, mais davantage de choisir les attributs pertinent en fonction de la position, et de tirer profit des paramètres discriminants suivant l'attitude de la cible.

En phase de reconnaissance, on retrouve le principe d'incrémentalité dans la construction de la solution, déjà décrit dans les chapitres précédents, mais cette fois, il est étendu à la sélection des attributs, des modèles. Nous préconisons de

[1] Ces modèles fournis par l'Aérospatiale ont été générés par le logiciel SPECRAY.

sélectionner des attributs, en fonction de l'hypothèse courante relative au type de la cible et à son attitude, de façon à découvrir son identité. La recherche de l'attribut le plus approprié permet de déterminer quelle est la donnée qui apporterait le plus d'informations. On garde la philosophie "anytime" qui précise qu'une solution est constamment possible, mais qu'elle est affinée avec le temps, et l'accumulation des connaissances.

D'un point de vue fonctionnel, cette approche nécessite la mise en œuvre de deux types d'agents que nous appellerons agent Identité et agent Position. Le schéma de principe est présenté sur la figure 6.2.

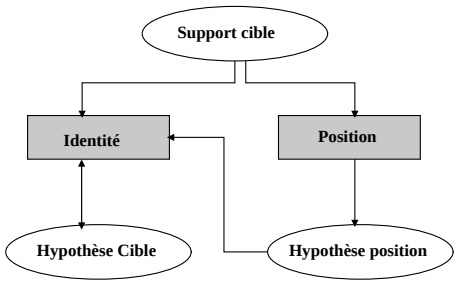


Fig. 6.2 – Description schématique du système de reconnaissance.

Le but est d'utiliser l'information position pour faciliter la reconnaissance. En fait, l'hypothèse de position validera des attributs supplémentaires de façon à prendre une décision sur le type de la cible. Le choix de ces attributs se fera de façon incrémentale et suivant un critère d'utilité.

Nous décrirons successivement dans ce chapitre les comportements des agents, les informations manipulées et le processus de décision.

6.3 Description des agents

Les deux nouveaux agents nécessaires à la mise en œuvre de la phase de reconnaissance sont les agents Position et Identité.

6.3.1 L'agent Position

Le rôle de l'agent Position est de définir l'attitude de la cible par rapport au capteur. A priori, on distingue trois types de vue possibles :

- les vues de Face/Dos ;
- les vues de Profil ;
- et les vues Dessus/Dessous.

Nous détaillerons cette catégorisation ultérieurement, dans le paragraphe 6.5.2.

L'agent Position utilise un ensemble de n attributs pour venir qualifier le type de vue. Donc à chaque support d'objet S, on va associer un vecteur d'attributs v[n]. De même que, pour la phase de détection, une mesure de distance entre ce vecteur et les modèles qualifiera l'hypothèse de position (HP).

$$(S_i, v[n]) \xrightarrow{P_C} HP_i$$

avec $HP_i \in \{Face, Profil, Dessus\}$

Cet agent ne travaille pas de façon incrémentale, en revanche le choix des n attributs peut être conditionné par les hypothèses de position formulées à partir des informations contextuelles et de la préparation de mission.

6.3.2 L'agent Identité

Le but de l'agent Identité est de conforter l'hypothèse de reconnaissance et d'identifier la cible. Est-ce un char, un navire, un avion ou un camion ? Son rôle est d'exploiter le support de l'objet mais avec un ensemble de p attributs plus spécifiques que les n précédents (p non inclus dans n), choisis éventuellement en fonction de l'hypothèse de position, lorsque celle-ci est disponible. Sachant une position, une distance entre le vecteur d'attribut v[p] et le modèle de cible (MC) donnera une hypothèse de cibles (HC).

$$(v[p], S_i, HP_i) \xrightarrow{D_H} HC_i$$

avec $i \in \{Char, Camion, Avion, Navire\}$

Ces agents ont des comportements qui sont basés sur la sélection d'attributs, suivant un critère d'utilité que nous décrirons ultérieurement. Mais au préalable, nous allons décrire la connaissance utilisée par ces agents à savoir les attributs eux-mêmes.

6.4 Les attributs pour la reconnaissance

En gardant la philosophie préconisée dans le chapitre relatif à l'état de l'art de la fonction détection, reconnaissance et identification, nous avons cherché une modélisation qui regroupe les informations géométriques, physiques et statistiques. Le modèle est donné sous forme d'un vecteur d'attributs.

Pour réaliser une première étude, nous avons choisi un grand nombre d'attributs de façon à évaluer leur utilité, lors d'une phase d'apprentissage. Nous aborderons ce problème dans le prochain paragraphe. Pour le moment, nous allons nous intéresser à la présentation de ces attributs.

Dans le cadre de notre application, nous avons principalement utilisé la caractérisation région pour modéliser la cible. Une approche contour donne une représentation sans doute plus fine, mais elle est plus difficile à extraire sur des images réelles complexes.

Les attributs sont mesurés soit sur les primitives précédemment extraites (support de cible), soit sur l'image brute. Nous distinguerons les attributs géométriques, physiques et statistiques.

6.4.1 Mesures géométriques

Mesures topologiques

Nous distinguons différentes mesures élémentaires :

- la surface est égale au nombre de pixels situés à l'intérieur de la frontière de l'objet.

- le périmètre est égal au nombre de pixels situés le long de la frontière de l'objet. Plusieurs méthodes peuvent être utilisées pour calculer le périmètre. À partir de la primitive région, nous avons choisi une approche de type amincissement, avec un élément structurant de taille 3*3 initialisé à 1. On incrémente le périmètre chaque fois que l'élément structurant ne peut pas être contenu dans la forme.

Ces attributs seront calculés à partir du support de cible et à partir de son enveloppe convexe. L'intérêt de ces attributs sera à terme d'exploiter des mesures de forme.

Il existe de nombreuses méthodes de construction d'une enveloppe convexe. Nous avons choisi d'implémenter l'approche de Graham [Preparara et Shamos, 1985]. Cet algorithme est détaillé en annexe C.3.

À l'issue de cet algorithme, l'ensemble des points sélectionnés décrivent une région sur laquelle il sera possible de calculer la surface (Senv) et le périmètre (Penv). Ces attributs seront ensuite exploités pour extraire des mesures de forme.

Ces mesures présentent l'avantage d'être toujours accessibles, simples et faciles à extraire. En revanche, elles sont fonction de la distance au capteur que l'on suppose toujours fixe.

Mesures d'orientation

Pour toutes les régions présentant des allongements caractéristiques, il est possible de calculer les directions privilégiées. Ces directions sont calculées à partir des moments d'inertie centrés du deuxième ordre. Elles sont décrites en annexe C.4.

Mesures de forme

Différentes méthodes sont utilisées pour caractériser la forme d'une région donnée. Nous présentons ci-dessous les six mesures utilisées :

- $C_1 = \frac{mp_1}{mp_2}$, mp1 et mp2 représentent les longueurs associées aux deux axes principaux. C1 permet de caractériser la compacité de la région,

- $C2 = \frac{S}{P^2}$, S et P représentent respectivement la surface et le périmètre de la région considérée. En théorie, cette valeur est maximale pour les cercles et elle est plus petite pour les autres formes. Toutefois, en raison de la discrétisation, cette valeur est maximale pour les polygones réguliers tels que les carrés et rectangles ou les polygones à peu près réguliers.

- $C3 = \frac{S}{Senv}$ rapport entre la surface et la surface de l'enveloppe convexe.

- $C4 = \frac{P}{Penv}$ rapport entre le périmètre et le périmètre de l'enveloppe convexe.

- $C_5 = \frac{P - Penv}{P}$ différence entre le périmètre et le périmètre de l'enveloppe convexe, normalisée par le périmètre.

- $C_6 = \frac{S - Senv}{S}$ différence entre la surface et la surface de l'enveloppe convexe, normalisée par la surface.

Ces mesures de formes permettent notamment d'évaluer la compacité du support qui pourra être un critère discriminatif entre les différentes cibles ou entre les différentes positions (figure 6.1).

Mesure de symétrie

Un critère de symétrie a été implémenté. Il est calculé de la façon suivante.

On considère successivement chaque point de l'axe principal d'inertie. Pour le premier point on compte dans la direction orthogonale le nombre de points appartenant au support et situés à droite de l'axe d'inertie, puis on retranche le nombre de point qui se situent à gauche. Cette différence définit la grandeur ϵ_1. Pour chaque point i de l'axe, on calcule ce ϵ_i.

La mesure de symétrie est ensuite définie par :

$$S = \sum_i \epsilon_i^2$$

L'idée d'utiliser un critère de symétrie paraît intéressante. Intuitivement, il semble utile dans la détermination de la position de la cible par rapport au capteur (par exemple entre un navire vu de profil et un navire vu de dessus). En général, les observations de dessus montrent une symétrie axiale. Toutefois, cette mesure a donné jusqu'à présent des résultats relativement décevants.

Les mesures d'angularité

Le coin est un point d'intérêt permettant de caractériser une cible. Le nombre de coins présent sur une région donnée permet de caractériser l'irrégularité de son contour.

Les détecteurs de points significatifs sont nombreux. Nos avons choisi d'utiliser le détecteur de Harris utilisé dans les travaux de C. Schmid [Schmid et Mohr, 1996]. Cet opérateur de Harris, présenté en annexe C.5, utilise uniquement les dérivées premières du signal, il évite ainsi une trop grande sensibilité au bruit.

6.4.2 Mesures physiques cinématiques

L'analyse dynamique des cibles permettra également d'extraire une information sur le type de la cible qui peut être très discriminante. Un mouvement important peut être caractéristique d'un avion. Le but est d'associer au centre de gravité de chaque support une mesure de vitesse et une mesure d'accélération et non plus seulement de réaliser une segmentation au sens du mouvement comme nous l'avions présenté dans la phase de détection. Cette mesure n'a pas encore été intégrée à notre système d'analyse.

6.4.3 Mesures statistiques

L'idée est toujours de chercher des attributs discriminants, mais cette fois en utilisant la distribution statistique des données.

Mesures classique

La texture traduit la valeur ainsi que la répartition des niveaux de gris pour une région donnée. Les méthodes permettant de quantifier ces mesures sont nombreuses. On utilisera principalement :

- La moyenne
- la variance
- l'écart type
- l'entropie

Ondelettes de Gabor

L'analyse fréquentielle va permettre une étude de la périodicité du signal. Une solution pour analyser le signal dans l'espace des fréquences est de faire une analyse

de Fourier. Le problème est que cette méthode n'est pas adaptée aux signaux ayant une fréquence variable au cours du temps. Le principe de calcul des ondelettes de Gabor est présenté en annexe C.6.

Conclusion

On dispose d'un ensemble d'attributs caractérisé par la diversité des informations considérées. L'aspect géométrique est traduit au travers des données de surface, de périmètre et de compacité. Les informations physiques sont représentées par la cinématique du support. Et enfin les informations statistiques sont caractérisées par la moyenne, la variance, l'entropie ou encore les coefficients en ondelettes. Disposant de tous ces attributs, l'objectif est désormais de les exploiter au mieux suivant un principe d'utilité, qui impose une phase d'apprentissage.

6.5 Apprentissage

Les informations regroupées dans un vecteur d'attributs caractérisant un support sont complexes. Ceux-ci peuvent être corrélés, entachés de bruit, non significatifs, etc. L'objectif est donc d'observer sur un ensemble d'apprentissage les attributs les plus pertinents pour trouver la position et l'identité d'une cible. L'outil utilisé a été l'analyse en composantes principales. Nous présenterons une ébauche de résultats qui demandera à terme une analyse plus approfondie.

6.5.1 Les images

La base d'apprentissage est constituée d'images de synthèse créées à partir du logiciel SPECRAY. Pour chaque type de cibles, nous avons généré des images dans 500 positions distinctes, et dans deux longueurs d'ondes différentes : visible et infrarouge. La caméra est positionnée autour de la cible suivant un quadrillage du type "ballon de football". Nous avons sélectionné les positions ayant une coordonnée positive en z, ce qui revient à choisir la demi-sphère supérieure. Six paramètres permettent de définir précisément la position de la caméra :

- Le roulis : oscillation latérale ;
- Le cap : la direction d'observation ;
- Le tangage : oscillation d'avant en arrière ;
- Les positions en x, y et z.

Les images initiales représentent différents types de cibles (chars, avions, véhicules, navires et hélicoptères). Un exemple d'image est présenté sur la figure 6.3. Chaque cible d'une taille d'environ 40*40 pixels, est présentée centrée sur un fond non bruité et non uniforme.

FIG. 6.3 – Image de synthèse générée par le logiciel SPECRAY.

6.5.2 Structuration de la base de données

Dans un premier temps, l'idée est de constituer une base d'apprentissage indexée par la position. Aussi, pour chaque cible et chaque longueur d'onde, nous allons définir des groupes de positions dans l'espace considéré. Le référentiel utilisé est décrit sur la figure 6.4. La cible est positionnée suivant l'axe des y.

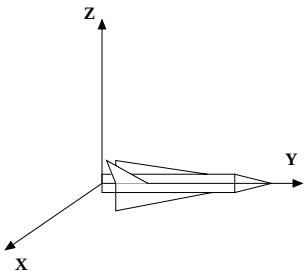


FIG. 6.4 – Le référentiel choisi.

Dans cet espace nous avons défini trois groupes de positions :

- De face, de derrière ;
- De profil ;
- De dessus, de dessous.

Des considérations de symétrie ont permis d'associer certains types de vue tels que face et derrière. En se replaçant dans le référentiel, on définit les cônes d'observation utilisés de la façon suivante :

Les axes des cônes sont définis par les médianes : $X = Y$, $Y = Z$, $X = Z$. La majorité des images fournies correspondent à des vues de dessus. Sinon, on dispose de quelques dizaines de vues de profil, et enfin seulement d'une vingtaine de vues de face.

Cette base de données, ainsi structurée, a été analysée. Le premier traitement a été d'extraire le support de la cible. Celui-ci a été effectué en deux temps. Un seuillage automatique des niveaux de gris a été réalisé, suivi d'un étiquetage en 8-connexité.

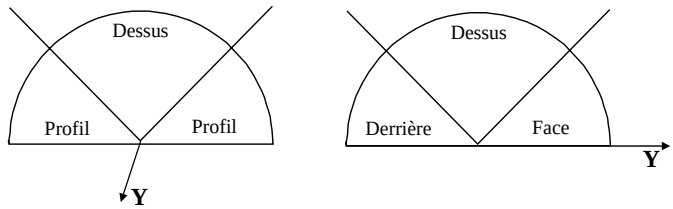


FIG. 6.5 – L'espace d'observation est composé de groupes de positions.

À partir des supports ainsi extraits, l'objectif a été finalement de regarder l'intérêt des différents attributs, au travers d'un procédé de fouille de données.

6.5.3 L'analyse en composante principale

L'analyse en composante principale est un outil pour trouver l'espace de représentation qui permet la meilleure discrimination des nuages de points. Le but est d'analyser les variations des attributs et leurs corrélations ainsi que la séparabilité des problèmes. Cette méthode descriptive consiste à résumer l'information en remplaçant les variables initiales par un nombre plus petit de variables composées linéairement. Elle permet notamment de faire une sélection entre les attributs significatifs et les attributs non significatifs. Elle souligne les fortes corrélations observées et élimine les fortes redondances. Pour visualiser les résultats de l'ACP, on projette les données sur les deux premiers axes principaux, dans l'espace transformé.

Pour la réalisation de cette expérience, nous avons développé une interface graphique, qui nous permet de sélectionner dynamiquement les attributs, et les groupes de cibles. Cette interface est présentée sur la figure 6.6.

La partie supérieure de cette interface permet de sélectionner des types de cibles représentés par des couleurs différentes (vert, marron, bleu), et des groupes de position distingués par des dégradés de couleurs. La partie inférieure gauche autorise la sélection d'attributs. On retrouve l'ensemble des attributs précédemment décrits. Enfin à droite, on active l'analyse en composante principale.

Voici quelques résultats obtenus par l'analyse en composantes principales. Le but est de reconnaître une cible étant donnée une position. La question sous-jacente est de savoir s'il est possible de trouver un vecteur caractéristique qui permette de reconnaître les différents types de cibles, connaissant leur position.

Plusieurs tests ont été réalisés. À chaque fois 8 paramètres ont été utilisés : la surface et le périmètre, la surface et le périmètre de l'enveloppe convexe, le rapport des périmètres, le rapport des surfaces, et deux critères de compacité $C2 = \frac{Penv^2}{Senv}$ et $C3 = \frac{mp_1}{mp_2}$. Voici les résultats obtenus :

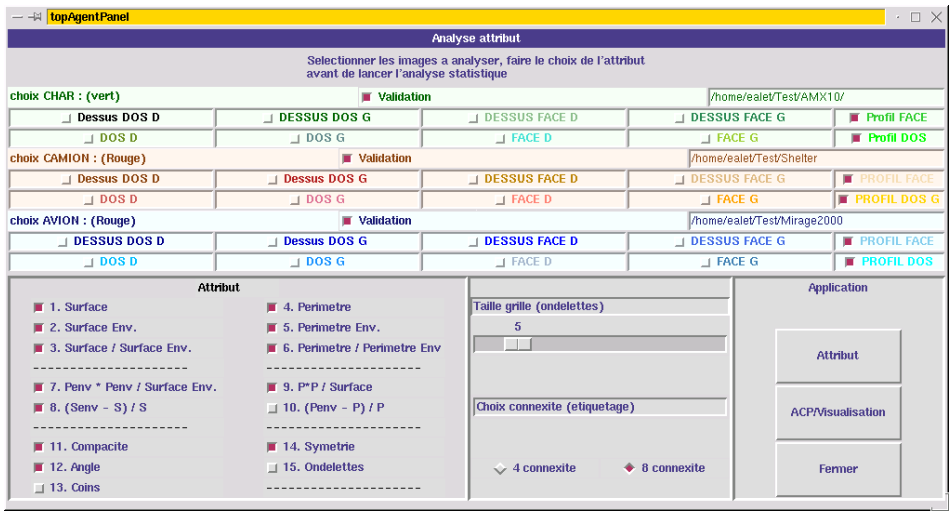


FIG. 6.6 – Interface pour réaliser l'analyse en composantes principales qui permet de sélectionner les attributs et les groupes de cibles.

– La première expérience consiste à analyser la répartition des nuages de points pour un type de cible et plusieurs groupes de positions. Sur la figure 6.7, on observe la distribution obtenue pour un char. En jaune-vert sont représentées les vues de face, en vert clair les vues de profil et enfin en vert foncé les vues de dessus.

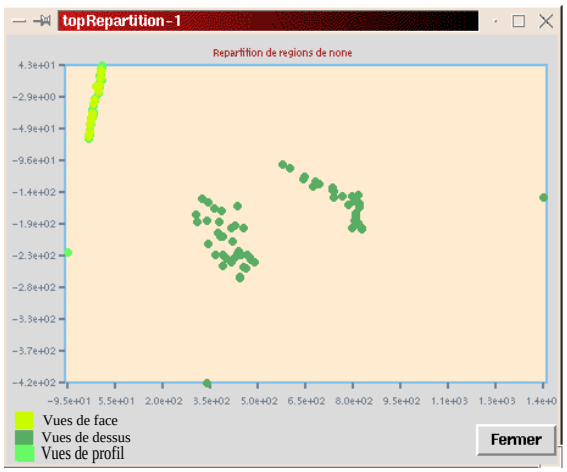


FIG. 6.7 – Analyse en composantes principales obtenue pour les chars vus de face(gris), de profil (en rose) et de face (rouge)

L'exemple du char est sans doute le cas le plus difficile a analyser, puisqu'il reste très compact quelque soit l'angle d'observation. Il est à ce sujet intéressant de voir que l'on arrive quant même à distinguer les vues de dessus. En revanche les observations de profil et de face restent confondues. Dans ce cas, relativement extrême, on peut quand même accèder à l'information

position, sachant que celle-ci peut ensuite être affinée par les informations contextuelles ou le type de mission. Forts de ce résultat nous nous sommes alors intéressés à la possibilité de distinguer les cibles connaissant une position.

– La seconde expérience consiste alors à regarder pour une position donnée, la répartition des nuages de points pour trois cibles différentes que sont les avions, les chars et les camions. Le résultat est présenté sur la figure 6.8. Les avions sont représentés en bleu, les chars en vert et enfin les camions en rose.

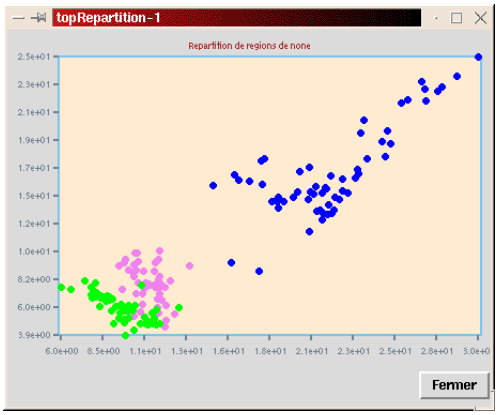


Fig. 6.8 – Analyse en composantes principales avec huit attributs pour trois types de cibles vues de profil.

Les attributs utilisés permettent de distinguer les trois nuages de points relativement distinctement. Avec les statistiques obtenues, nous avons pu constater que le critère de compacité était dans ce cas prépondérant.

Nous avons renouvellé cette expérience avec des observations de face des différentes cibles. Le résultat obtenu est présenté sur la figure 6.9. Le code de couleur a été conservé. Les avions sont représentés en bleu, les chars en vert et les camions en rose.
Les huit attributs utilisés permettent uniquement de distinguer les avions et les véhicules au sol . Les chars et les camions sont confondus pour les vues de face. Le critère de compacité n'est plus suffisant pour les distinguer.

Avec un ensemble d'attributs restreints, il est possible de distinguer le plus souvent la cible pour une position donnée. Les avions se distinguent facilement, en revanche il existe encore une difficulté pour différencier les chars et les camions. Toutefois ces deux expérimentations justifient déjà l'intérêt de l'agent Position. Puisque, lorsque l'information position est disponible, la reconnaissance de la cible peut être facilitée.

Maintenant entre ces constations et la notion d'utilité, il existe encore une difficulté. Nous proposons de faire une analyse factorielle discriminante, de façon à traduire les résultats sous forme probabiliste.

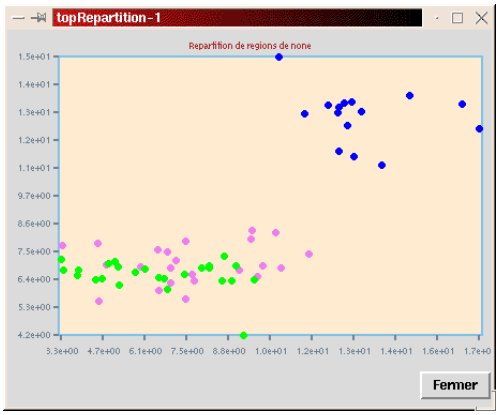


FIG. 6.9 – Analyse en composantes principales avec huit attributs pour trois types de cibles vues de face.

6.5.4 L'Analyse factorielle discriminante (AFD)

L'analyse factorielle discriminante est un cas particulier de L'ACP. Faire une AFD est équivalent à faire une ACP sur le nuage des centres de gravité.

Notre but

Pourquoi faire une analyse discriminante ? C'est la première question que l'on est en droit de se poser. Le but est d'exploiter notre connaissance sur les attributs pour ensuite initialiser la connaissance dans le système par l'intermédiaire de réseaux bayésiens que nous décrirons dans la partie suivante. L'analyse discriminante permet de construire à partir d'un ensemble de variables, des fonctions discriminantes qui définissent une règle de décision optimale à partir de laquelle on peut affecter des individus tests aux différents groupes.

L'approche probabiliste

Nous supposons ici connaître la distribution $f_i(x)$ d'avoir un individu x dans le groupe C_i. La probabilité d'assigner x au groupe C_i sachant qu'il appartient au groupe C_j est notée P[i/j].

Définition de la règle d'affectation :

Il y a d mesures pour chaque individu. C'est-à-dire qu'il correspond à chacun un vecteur d'attributs dans l'espace R^d. Maintenant divisons l'espace R^d en deux régions mutuellement exclusives R_1^d et R_2^d (dans le cas d'un problème à deux classes).

La probabilité d'erreur est donnée par :

$$P[1/2] = \int_{R_1^d} \int \dots \int f_2(x)$$

$$P[2/1] = \int_{R_2^d} \int \dots \int f_1(x)$$

Ceci impose la règle d'affectation suivante :

$$\frac{f_1(x)}{f_2(x)} > \frac{\pi_2}{\pi_1}$$

avec π_i la population d'une classe C_i

Soient C(1/2) et C(2/1) les coûts de mauvaise classification, alors le coût total s'écrit :

$$\frac{f_1(x)}{f_2(x)} > \frac{\pi_2.C(1/2)}{\pi_1.C(2/1)}$$

Ce problème est un cas particulier en théorie de la décision et cette relation est en fait une relation bayésienne. L'intérêt est d'avoir accès à la grandeur "coût".

Cette technique n'a pas été développée, mais déjà elle permet de faire le lien avec la phase de détection et l'utilisation des réseaux bayésiens.

6.6 La fonction utilité

Sachant que les fausses alarmes aussi bien que les non détections ont un coût, l'objectif est bien de minimiser le coût global. Dans la phase de détection, la fonction coût était exprimée au travers des réseaux bayésiens. Il s'avère qu'une version spécifique de ces réseaux, appelée *diagramme d'influence*, permet de les adapter à ce type de problème [Becker et Naïm, 1999]. Ces diagrammes donnent une représentation graphique du problème de la décision [Howard et Matheson, 1984].

6.6.1 Diagramme d'influence

Les diagrammes d'influences sont des réseaux bayésiens auxquels on a adjoint des variables de décision et une fonction d'utilité [Pearl, 1988]. Ces variables sont représentées par des noeuds dans le graphe :

- les noeuds de hasard sont désignés par des cercles. Ils représentent les variables aléatoires du système.
- les noeuds de décision sont désignés par des carrés. Ils représentent le choix d'une décision ;
- les noeuds d'utilité sont figurés par des losanges. Ils représentent l'objectif (l'utilité) à maximiser.

En toute rigueur, dans un diagramme d'influence, il ne devrait y avoir qu'un seul noeud d'utilité puisque cette utilité est représentée par un scalaire. Cependant, on peut faire l'hypothèse que la fonction utilité se décompose en une somme de fonctions d'utilité élémentaires, d'où la possibilité d'avoir plusieurs noeuds d'utilité.

En ce qui concerne les arcs des graphes, selon leurs destinations, leur signification varie.

- les arcs pointant vers les nœuds d'utilité ou de hasard représentent une dépendance probabiliste ou fonctionnelle, et pas forcément une dépendance de causalité ou d'antériorité, comme c'est souvent le cas.

– les arcs pointant vers les nœuds de décision, n'ont qu'une valeur informative. Ils précisent les variables aléatoires qui seront considérées pour prendre la décision.

Nous allons désormais appliquer cette représentation à notre problème et proposer un diagramme d'influence adapté.

6.6.2 Un exemple de diagramme

D. Suryadi et P.J. Gmytrasiewick [Suryadi et Gmytrasiewicz, 1999] proposent une approche dans le domaine militaire avec une application à la poursuite de cibles et à la défense anti-aérienne. L'exemple proposé est illustré sur la figure 6.10

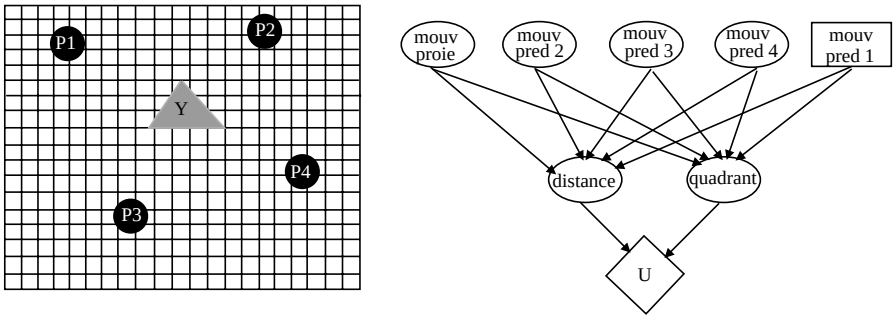


FIG. 6.10 – Exemple d'utilisation des diagrammes d'influence dans le domaine de la poursuite. À gauche, le domaine de poursuite ainsi que les acteurs du jeu sont présentés. À droite, le diagramme d'influence nécessaire à la décision du prédateur 1, est décrit.

La poursuite d'une proie Y est prise en charge par quatre agents P_1, P_2, P_3 et P_4 qui ont un rôle de prédateur. Ils entourent la proie dans l'environnement 2D de la grille. Ces agents bougent simultanément et le jeu s'arrête lorsque :

– tous les prédateurs ont vu la cible sous les quatre angles possibles (quatre quadrants).
– le temps est écoulé.
– la proie est poussée contre un mur (côté de la grille).

Pour capturer la cible, les agents doivent coordonner leurs actions. Sur la partie droite de la figure 6.10, le modèle de décision associé au prédateur numéro 1 est présenté. Ce schéma de contrôle intègre le mouvement des autres prédateur. Dans ce modèle, l'utilité est influencée par la distance de tous les prédateurs à la cible, et par le nombre de quadrants d'observation.

Nous pouvons envisager de reprendre ce schéma dans le cadre du processus de décision de l'agent Identité.

Considérons un exemple avec plusieurs variables aléatoires binaires :

– Atti : représente les attributs utilisés. $I \in [1, N]$ avec N le nombre total d'attributs ;

- P : représente le nœud Position à savoir l'attitude d'une cible par rapport au capteur. C'est un noeud à trois états : dessus, profil, dos ;
- un nœud D représente la décision de poursuivre l'accumulation d'informations avec la sélection de nouveaux attributs ;
- un nœud U d'utilité est une fonction de toutes les variables précédentes, qui représente le coût de la situation.

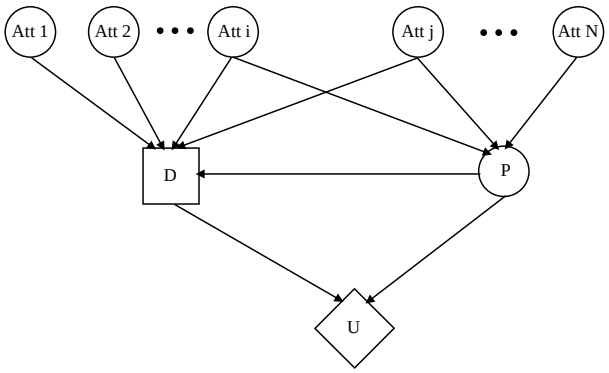


Fig. 6.11 – Exemple de diagramme d'influence modélisant un problème de décision consistant à sélectionner un attribut supplémentaire.

La position est une variable à trois états. Trois types de vue sont considérées : dessus, profil et face. D est un nœud de décision binaire, il est vrai si on décide d'ajouter un attribut supplémentaire et faux sinon. Le nœud U prend des valeurs réelles et représente l'utilité, grandeur que l'on cherchera à maximiser.

Dans le cadre des diagrammes d'influence il faut ajouter aux tables des probabilités conditionnelles, les tables d'utilité qui devront être extraites par apprentissage et notamment avec l'analyse en composantes principales.

6.7 Conclusion

Dans la description du système en phase de reconnaissance, nous avons introduit la notion d'attributs pour caractériser la position et la cible en question. À partir de là, deux problèmes se posent :

- le premier concerne l'utilité des attributs. Par exemple pour une vue donnée, il peut être intéressant d'utiliser tel ou tel attribut plutôt que tel autre. Il faut, par conséquent en fonction de la vue et du contexte, associer une mesure d'utilité à chaque attribut ;

- le second concerne la capitalisation de la connaissance, et la façon dont l'information va se transmettre dans le système. Comment va se propager une hypothèse de cible ? Nous avons choisi une approche probabiliste par l'in-

termédiaire des réseaux bayésiens et plus spécifiquement des diagrammes d'influence. Toutefois le réglage de cette structure impose un apprentissage pour régler les pondérations nécessaires.

Bien évidemment la sélection de variables introduit des contraintes supplémentaires. Il faut par exemple, réapprendre le réseau à chaque étape, avant d'estimer à nouveau la pertinence des variables.

Cette phase de reconnaissance n'a pas encore été intégrée dans le système mais elle nous a permis d'étendre la philosophie proposée pour la phase de détection tout en respectant les principes :

- de focalisation et d'adaptation : le système va porter son attention sur les attributs discriminants pour un type de cible et choisir les méthodes adéquates.
- de coopération et de distribution. La tache de vision est réalisée par la mise en œuvre de deux agents (Position et Identité).
- d'incrémentalité et d'apprentissage dans le choix des attributs
- d'incertitude avec cette fois une modification ; on propose une évolution des réseaux bayésiens vers une représentation de type diagramme d'influence.

Chapitre 7

Conclusion

L'objectif de ce travail était de proposer une architecture pour réaliser la fonction détection, reconnaissance et identification de cibles militaires dans des contextes variés. La détection consiste à déterminer la présence éventuelle d'une cible, la reconnaissance vise à lui associer une classe et enfin l'identification permet de la nommer. Ce travail s'inscrit dans un contexte variable (mission air-sol, sol-air, air-air, sol-sol) et multi-capteur (visible et infrarouge). Nos connaissances *a priori* sur les éléments de scène et les cibles sont représentées au sein de modèles acquis de façon interactive (interface d'acquisition). Les modèles diffèrent suivant :

- les informations manipulées : géométriques, statistiques, physiques ou encore symboliques ;
- la nature de la caractérisation : locale ou globale .

Nous avons choisi d'intégrer l'ensemble de ces critères pour modéliser les éléments d'une scène. Les différentes informations permettent d'assurer une richesse dans la caractérisation. Les données symboliques sont intégrées très tôt dans le processus DRI afin de confirmer notamment des hypothèses courantes. Les caractérisations globale et locale ont permis d'assurer respectivement une robustesse vis-à-vis du bruit et des occultations éventuelles.

La mise en correspondance entre le modèle *a priori* et le modèle d'observation se fait au travers d'un calcul de distance. Elle permet ainsi d'évaluer un coefficient d'appartenance à un concept donné. Cette incertitude est propagée au sein du système grâce à des réseaux bayésiens.

Ces choix permettent de répondre uniquement à certaines difficultés. En effet, en plus de la robustesse aux perturbations, de la grande quantité d'informations à gérer, et de l'hétérogénéité des connaissances, nous devons faire face à une très grande variabilité interclasse et intraclasse. Ces contraintes nous ont donc conduit vers une architecture multi-agent capable de distribuer la connaissance et d'effectuer

des traitements en parallèle et en concurrence.

Nous avons posé le problème de la DRI comme un problème ancré dans un espace à trois dimensions (l'espace image, l'espace des buts et l'espace des méthodes). Cette articulation est présente dans la base de connaissances avec les liens quoi-comment et quoi-où, mais elle s'intègre également dans les principes fondamentaux de notre système au travers de :

- la focalisation et l'adaptation ;
- la distribution et la coopération ;
- la gestion de l'incertitude ;
- l'incrémentalité et l'apprentissage.

Chaque principe permet au système de naviguer au sein d'un espace, mais également de basculer d'un espace à l'autre. Nous avons notamment insisté sur la façon dont les principes de focalisation et d'adaptativité se déclinent au sein de ces espaces de représentation.

Ces principes sont également étroitement liés entre eux. Ils s'articulent les uns par rapport aux autres. Par exemple, la focalisation est la première étape avant de penser à l'adaptation. De même la distribution est un préalable au principe de coopération. Les résultats expérimentaux ont permis de souligner d'autres liens comme par exemple le lien confiance-stratégie. Pour différentes stratégies proposées, des confiances différentes seront obtenues et inversement suivant les confiances calculées, les stratégies mises en œuvre seront différentes.

Le système proposé élabore sa propre stratégie suivant les informations disponibles. Il procède de façon incrémentale. Il émet des hypothèses qui sont confirmées ou non, par la suite, en fonction des informations collectées. L'approche proposée suit une philosophie "anytime" c'est-à-dire qu'il existe toujours une décision possible, mais celle-ci sera affinée avec temps.

Notre principal apport a été la définition d'une méthodologie générale pour un système DRI. La démarche a été soigneusement détaillée pour la phase détection, et très facilement étendue à la phase de reconnaissance et d'identification. Mais, toujours pour affiner les qualités d'adaptation du système, nous proposons d'améliorer les choix de stratégies en basant notre raisonnement sur une fonction coût-utilité. L'objectif est de déterminer l'ensemble d'attributs qui autorisera la meilleure discrimination possible. Aussi, l'intégration de noeuds d'utilité au sein des réseaux bayésiens est une future étape.

Les qualités du système sont indéniables, même si des améliorations sont encore possibles. Le point clé reste sans doute l'apprentissage qu'il faut envisager à plusieurs niveaux :

- dans la modélisation des objets et la spécification des règles de focalisation afin de remplacer l'outil interactif qui a des limites certaines.
- dans la conception et dans la paramétrisation des réseaux bayésiens.

Cet apprentissage permettra encore une plus grande indépendance et autonomie du système. Dans la base de connaissances, un apprentissage par renforcement permettrait à la fois de conserver la philosophie d'accumulation incrémentale de la connaissance, tout en se détachant d'un *a priori* parfois approximatif ou non représentatif du contexte applicatif.

Annexe A

Les réseaux bayésiens : “The bayes Net Toolbox (BNT)”

Nous avons choisi d'utiliser la “Bayes Net Toolbox” développée par K. Murphy pour intégrer l'approche des réseaux bayesiens dans le système de détection et de reconnaissance de cibles. Cet outil, développé dans l'environnement Matlab, permet de modéliser et d'instancier différents modèles graphiques tels que les réseaux bayesiens, les modèles de Markov cachés... Une documentation complète est proposée à l'adresse suivante :

$$http://www.cs.berkeley.edu/\%7Emurphyk/Bayes/bnt.html.$$

Une version en français a été proposée par P. Leray à l'adresse :

$$http://asi.insa-rouen.fr/enseignement/siteUV/rna/BNT$$

Notre objectif ici n'est donc pas de refaire une documentation complète, mais davantage de donner les points clés dans la construction d'un réseau, en se basant sur un exemple simple.

A.1 Exemple de réseau

Le principe de construction des réseaux bayésiens se décompose en trois étapes. Il faut :

- recenser toutes les variables aléatoires utilisées ;
- choisir un ordre pour ces variables, à savoir les causes puis les conséquences ;
- associer pour chaque noeuds sa probabilité conditionnellement à ses parents ;
- puis calculer les probabilités recherchées.

Nous allons illustrer la mise en œuvre d'un réseau en analysant la démarche effectuée pour un exemple simple proposé initialement par K. Murphy.

Exemple : ce matin là, alors que le temps est couvert, M. X constate que sa pelouse est humide. Il se demande alors s'il a plu pendant la nuit, ou s'il a simplement oublié de débrancher son arroseur automatique !

Le graphe associé à cette situation est composé de quatre noeuds représentant quatre variables aléatoires : le temps humide, l'arroseur, la pelouse humide et la pluie. Il est représenté sur la figure A.1. Les variables sont booléennes. Les états possibles sont “vrai” et “faux”.

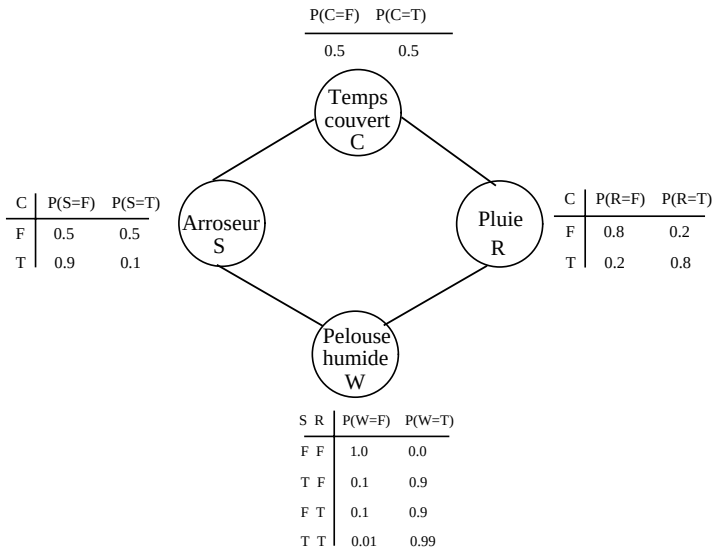


Fig. A.1 – Un exemple de réseau bayésien composé de 4 variables aléatoires.

À chaque noeud du graphe est associée une table de probabilités conditionnelles décrivant les liens de dépendance.

A.2 Construction d'un graphe

Pour construire le réseau, il faut commencer par donner le nombre de noeuds du graphe et initialiser à zéro une matrice d'adjacence de la taille correspondante.

```
N=4 ;
dag=zeros(N,N) ;
```

Les liens de causalité entre les noeuds sont ensuite spécifiés dans la matrice d'adjacence. Chaque noeud correspond à une ligne et à une colonne de la matrice. La numérotation adoptée est la suivante : C=1, S=2, R=3, W=4.

Donc les liens de causalité sont définis par :

```
dag(R,W)=1 ;   ceci signifie qu'il y a un arc entre les noeuds R et W.
dag(S,W)=1 ;
```

```
dag(C,[R,S])=1 ;
```

Concernant la description des noeuds, il faut ensuite spécifier leur type et leur taille c'est-à-dire le nombre de valeurs possibles dans le cas de noeuds discrets. Nous avons fait l'hypothèse de variables binaires, aussi :

```
discrete_nodes = 1 :N ;
taille_noeud=2*ones(1,N) ;
```

Maintenant que les arcs et les noeuds sont décrits, il suffit alors de créer le réseau avec la commande suivante :

```
bayesnet=mk_bnet(dag,taille_noeud)
```

Dans cette structure de graphe, il est nécessaire de préciser les paramètres du modèle. Nous devons spécifier les probabilités *a priori*, et plus précisément la loi de distribution des probabilités conditionnelles (CPD = conditional probability distribution) relative à chaque noeud. Dans le cas des variables discrètes, les probabilités conditionnelles sont représentées sous forme de tableaux indexés par les valeurs des parents éventuels. La syntaxe utilisée est la suivante :

```
bnet.CPD{C} = tabular_CPD(bnet, C, [0.5 0.5])
bnet.CPD{R} = tabular_CPD(bnet, R, [0.8 0.2 0.2 0.8])
bnet.CPD{S} = tabular_CPD(bnet, S, [0.5 0.9 0.5 0.1])
bnet.CPD{W} = tabular_CPD(bnet, W, [1 0.1 0.1 0.01 0 0.9 0.9 0.99])
```

À ce stade, le réseau est entièrement défini. On connaît les noeuds, les arcs et les probabilités conditionnelles associées à chaque noeud.

A.3 L'inférence

Il existe différents algorithmes pour mettre en oeuvre l'inférence dans les réseaux bayésiens. Ils diffèrent notamment du point de vue de la rapidité, de la complexité, de la précision. Nous étudierons ici uniquement la structure de type "arbre", utilisée dans le cadre des variables discrètes. L'inférence est spécifiée par la commande :

```
engine = jtree_inf_engine(bnet) ;
```

Par exemple, si nous voulons calculer la probabilité qu'il a plu, sachant que l'herbe est mouillée P(W=True) et que les autres noeuds ne sont pas précisés, nous devons écrire :

```
evidence = cell(1,N) ;
evidence{W} = true ;
```

La représentation des évidences se fait grâce à un tableau de cellules à une dimension. Elle permet une représentation des noeuds sous forme de vecteurs de longueurs différentes. Nous pouvons désormais ajouter l'observation à notre structure :

```
[engine, ll] = enter_evidence(engine, evidence);
```

Le premier argument retourne la structure modifiée et le second argument retourne la log-vraissemblance. Finalement, il est possible de calculer la probabilité P(R=True/W=True) :

```
m = marginal_nodes(engine, R);
p1 = m.T(true);
P(R=true|W=true) = 0.4298
```

A.4 Vérification par le calcul

$$P(R = True/W = true) = \frac{P(R = True, W = true)}{P(W = 1)}$$
$$P(R = True/W = true) = \frac{P(C = c_1, S = s_1, R = true, W = true)}{P(C = c_1, S = s_1, R = r_1, W = w_1)}$$

$$\begin{aligned}
P(C = c_1, S = s_1, R, W) &= P(C = c_1)P(S = s_1/C = c_1).P(R = 1/C = c_1).P(W = 1/R = 1, S = s_1) \\
P(C = c_1, S = s_1, R, W) &= P(C = 1)P(S = 1/C = 1).P(R = 1/C = 1).P(W = 1/S = 1, R = 1) \\
&+ P(C = 1)P(S = 0/C = 1).P(R = 1/C = 1).P(W = 1/S = 0, R = 1) \\
&+ P(C = 0)P(S = 1/C = 0).P(R = 1/C = 0).P(W = 1/S = 1, R = 1) \\
&+ P(C = 0)P(S = 0/C = 0).P(R = 1/C = 0).P(W = 1/S = 0, R = 1) \\
P(C = c_1, S = s_1, R, W) &= (0.5 * 0.1 * 0.8 * 0.99) + (0.5 * 0.9 * 0.8 * 0.9) \\
&+ (0.5 * 0.1 * 0.2 * 0.99) + (0.5 * 0.9 * 0.2 * 0.9) \\
P(C = c_1, S = s_1, R, W) &= 0.4545
\end{aligned}$$

Suivant un calcul similtaire, on trouve :

$$P(C = c_1, S = s_1, R = r_1, W = w_1) = 0.6471$$

Donc :

$$\begin{aligned}
P(R = True/W = true) &= \frac{P(R = True, W = true)}{P(W = 1)} \\
&= \frac{0.4545}{0.6471}
\end{aligned}$$

$$\boxed{\text{P(R=True/W=true)=0.702}}$$

Annexe B

Le modèle d'agent, l'administrateur et les fiches d'informations

B.1 Le modèle d'agent

Nous donnons ici une description du modèle d'agent générique utilisé [Bouillot, 1999] :

```
class agent_c{
protected :
   //structure pour gestion de temps des messages
   struct timespec to ;

   // nom de la thread et identificateur de l'agent
   pthread_t tid ;
   int agent_id ;
   bool etat ;

   // variables liées à la gestion des threads
   pthread_mutex_t mtx_mail ;
   pthread_mutex_t mtx_etat ;
   pthread_cond_t cond ;

   // variables liées à la gestion des messages
   int *type_message ;
   int *type_contenu ;
   unsigned long *date_message ;
   agent_c* *interlocuteur ;
   void* contenu ;

   // variables et méthodes liées au mécanisme de boite aux lettres (BAL)
   int *pointeur_lettre ;
   BAL<message_c> b_lettre ;
   void Mailbox(message_c *courrier) ;
   void lock_mailbox() ;
```

```
   void unlock_mailbox();
public :
   agent_c(int aid);
   virtual  agent_c(void);
   ... métodes de lecture de messages
   ... méthodes d'envoi de messages ...
   };
```

B.2 L'administrateur

Il est composé par un ensemble de variables et de méthodes, que nous allons brièvement détailler

```
class admin_c {
private

   //Les fiches d'information
   BAL<info_agent_t> boite_a_fiche;

   //variables liées à la gestions des threads
   pthread_mutex_t mtx;
   pthread_mutex_t mtx_affi;
   pthread_mutex_t mtx_gestion;

   //variables liées au nombre d'agents
   int nb_agt_initiaux;
   int nb_agt_presents;
   int nb_agt_tues;

   //variables liées à la gestion du temps
   struct timeval t0,t;
   struct timezone *tzo;

   //variables liées à la tracabilité des messages
   char nom_fic_mouchard[32];
   FILE *fic_mouchard;

public :
   admin_c();     ~admin_c(void);
   void affi_info_agent(info_agent_t*,char*);
   void enregistrer(agent_c*,int);
   void deenregistrer(agent_c*);
   info_agent_t* identifier(const int);
   void activer();
   unsigned long horloge();
};
```

Les variables associées à l'objet administrateur se décomposent en cinq parties.

Il existe une variable **info_agent_t** qui capitalise les informations sur les agents du système. Nous détaillerons cette structure dans le paragraphe suivant. L'accès à

ces données peut se faire par l'appel de la fonction **affi_info_agent**

Un ensemble de variables liées à la gestion des mutex est également pris en compte (**mtx,mtx_affi, mtx_gestion**). Elles permettent de gérer le verrouillage et la protection des accès en mémoire. Elles seront utilisées notamment pour mettre à jour la liste des fiches d'informations.

Les structures **timeval** et **timezone** permettent de gérer la notion de temps. Elle permettent de fixer un repère temporel commun pour tout le système. L'accès à la donnée temps se fait par l'intermédiaire de la fonction **horloge()**.

Enfin, les variables **nom_fic_mouchard** et **fic_mouchard** ont une utilité dans la traçabilité au cours du temps. Le fichier fic_mouchard contient l'ensemble des messages délivrés par les agents.

L'ensemble des données ou des méthodes décrites jusqu'à présent permettent de stocker ou de visualiser des données. Il faut désormais s'intéresser aux mécanismes dynamiques liés à l'administrateur. Cette dynamique est gérée par l'intermédiaire des fonctions **enregistrer**, **activer** et **deenregistrer**. L'enregistrement se caractérise par la création d'une fiche d'informations ainsi que l'affectation d'une zone mémoire suffisante en fonction des besoins de l'agent. L'activation correspond à l'activation d'un thread pour chaque agent. Un agent est donc du point de vue informatique un processus autonome. Il a la possibilité de se détruire en faisant appel à la fonction deenregistrer. La liste des fiches d'informations est alors parcourue, et la fiche associée à l'agent en question est supprimée.

B.3 Les fiches d'informations

Cette information est composée de la façon suivante :

```
struct info_agent_t{
   agent_c *agent ;
   pthread_t tid ;
   int type ;
   char *nom ;
   Tableau_Representation *zone_memoire ;

   info_agent_t ( agent_c *pag,int type_agt,char *nom_agent, Tableau_Representation
*mem) { agent = pag ; tid=0 ; type = type_agt ; nom = nom_agt, zone_memoire = mem ;}
}
```

Dans la fiche d'information est mémorisée l'adresse de l'agent considéré (**agent**), l'identificateur de la thread associé (**tid**), le type de l'agent (**type**), son nom (**nom**) et enfin l'adresse de la zone mémoire qui lui est attribué (**zone_memoire**)

B.4 Communication entre agents

Chaque agent dispose d'une boîte aux lettres assimilable à un tableau dynamique. Cette structure permet une gestion des messages et plus précisément :

- le stockage des messages ;
- la destruction des messages ;

Chaque message comporte un intitulé, une date, un auteur, une information qualitative sur le contenu et enfin le contenu lui-même. Suivant ces paramètres, l'agent peut définir des priorités dans la lecture des messages.

Annexe C

Algorithmes de traitement d'images

C.1 Les méthodes et leurs paramètres

C.1.1 Les méthodes de l'agent Localisation

On distingue quatre méthodes différentes basées sur ;
- les niveaux de gris
- l'homogénéité
- la transformée de Hough
- la surface

Chacune de ces méthodes est décrite par une condition et un ensemble de paramètres.

Les niveaux de gris

La syntaxe utilisée pour cette méthode est :

gris : cond : min : max

L'objectif est de regarder dans un voisinage le nombre de pixels qui ont un niveau de gris compris entre "min" et "max". Si ce nombre est supérieur au paramètre "cond", le pixel courant est validé. Le voisinage est spécifié dans la base de connaissances avec la variable TAILLE.

L'homogénéité

La syntaxe utilisée pour cette méthode est :

homog : cond : min : max

De la même façon que précédemment, on regarde dans un voisinage les variations d'amplitude entre le pixel central et chacun de ses vosins. Dès que l'amplitude d'une transition est comprise entre "min" et "max", un compteur est incrémenté. Lorsque tout le voisinage a été parcouru, on compare le compteur au paramètre "cond". Si le compteur est supérieur ou égal à "cond", le pixel courant est validé

La transformée de Hough

La syntaxe utilisée pour cette méthode est :

hough : cond : param1 : param2

La syntaxe est un peu différente par rapport aux méthodes précédentes. Il y a deux paramètres nécessaires au calcul de la transformée de Hough. "Param1"et "param2" sont des seuils de validation respectivement utilisés sur les niveau de gris et dans la table d'aacumultation pour étiquetter les pixels. La condition de validation est effectuée dans l'espace de Hough. Toujours dans un voisinage si le nombre de pixel validé est supérieur ou égal à "cond", alors le pixel courant est validé.

La surface

La syntaxe utilisée pour cette méthode est :

surface : 0 : min : max

Dans ce cas, il n'y a pas de condition de validation (cond = 0), mais uniquement un intervalle de validation. En effet les trois méthodes précédentes permettent de définir des régions de l'image. Avec cette méthode, on filtre ces régions suivant un critère géométrique. Ne seront validées que les régions ayant un nombre de pixels compris entre "min"et "max".

C.1.2 Les méthodes des autres agents

Mis à part l'agent Localisation, pour les autres agents, une seule méthode a été implémentée.

L'agent détection

Il utilise une technique basée sur une croissance de régions. La syntaxe est la suivante :

détec : cond : 0 : 0

Cette méthode utilise seulement un paramètre pour la croissance qui permet de pondérer l'intervalle de validation défini suivant des statistiques locales.

L'agent focalisation

Il utilise une méthode de focalisation dont la syntaxe est

FOCUS : type : obj

Le premier paramètre "type", spécifie la nature de la focalisation (sur, à côté ou ailleurs dans l'image). Le paramètre obj spécifie le nouvel objectif.

L'agent mouvement

Il utilise une méthode de mouvement qui sera décrite dont la syntaxe est

MOUV : cond : min : max

L'image mouvement obtenu par l'algorithme est une image en niveaux de gris. Aussi suivant le même principe que pour la méthode de Localisation basée sur les niveaux de gris, on valide le pixel courant uniquement si un nombre suffisant de voisin ont un mouvement compris entre "min" et "max".

C.2 La transformée de Hough :

C.2.1 Le principe

La transformée de Hough est usuellement utilisée pour la détection de courbes analytiques telles que les lignes, les cercles, les ellipses... Elle permet une extraction de caractéristiques géométriques globales de l'image.

Le principe est de calculer en chaque point $P_i(x_i, y_i)$ du contour d'un objet l'ensemble des paramètres α qui vérifient :

$$\mathrm{H}(x_i,y_i,\alpha)=0$$

Prenons l'exemple d'un point dans le plan image, cherchons sa transformée dans l'espace des paramètres (voir la figure C.1) :

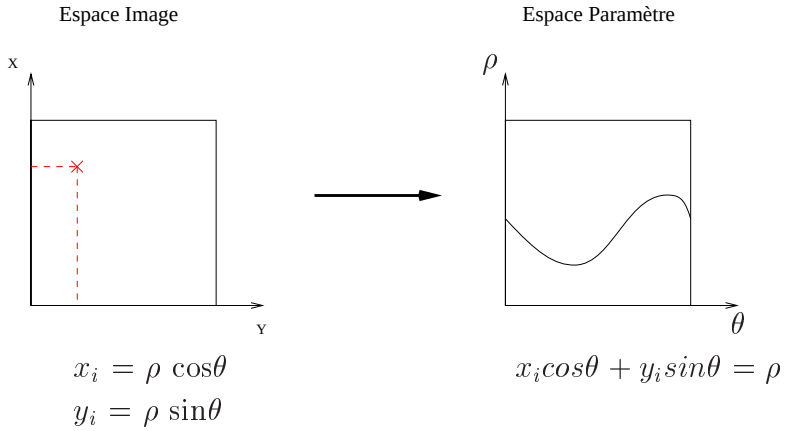


Fig. C.1 – La transformée de Hough d'un point

La transformée d'un point est une sinusoïde d'équation $x_i cos\theta + y_i sin\theta = \rho$. De ce résultat, nous pouvons en déduire la transformée d'une droite puisque celle-ci est définie par deux points. Elle est caractérisée par une distance ρ et un angle θ, fixés. Donc dans l'espace des paramètres (ρ,θ), la représentation sera un point (figure :C.2).

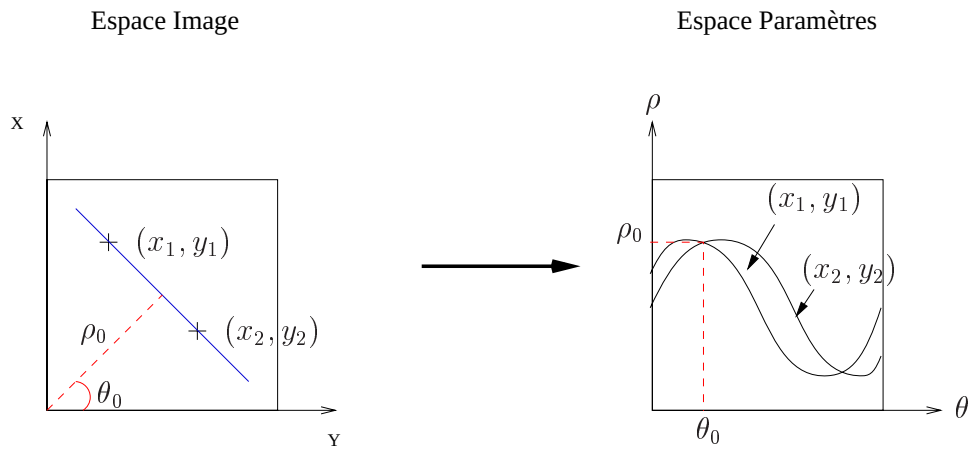


FIG. C.2 – La transformée de Hough d'une droite

En pratique, le calcul de la transformée de Hough d'une image binaire, se décompose en plusieurs étapes :

- initialisation d'un compteur 2D (accumulateur), de la taille de l'image, à zéro.
- chaque point de l'image originale, est représenté par une sinusoïde dans l'espace des paramètres. Et chaque fois que la sinusoïde traverse une case, le compteur est incrémenté. Ainsi, pour une intersection de deux courbes, le compteur passe à 2.
- le maximum global du compteur permet de définir un ρ et un θ, soit une droite dans l'espace image.
- lorsque le compteur est supérieur à un seuil prédéfini une droite est détectée

La transformée de Hough peut être généralisée à la détection de structures dans une image prétraitée, sachant que le prétraitement sélectionne les points caractéristiques sur lesquels travaillera la Transformée de Hough.

Un des avantages de cette méthode est sa robustesse au bruit et sa tolérance faces aux discontinuités. À partir de ce schéma de base concernant la transformée de Hough, différentes mises en oeuvre ont été proposées. Nous distinguerons notamment les approches déterministes, et les approches probabilistes. Nous présenterons ici les principales méthodes utilisées, ainsi que leurs avantages respectifs.

C.2.2 Les méthodes déterministes

La transformée de Hough Standard (SHT) :

Cette technique se décompose en trois phases successives :
- calcul des valeurs des paramètres et mise à jour de l'accumulateur pour tous les points de l'image,
- recherche des maxima locaux dans l'accumulateur,
- extraction des segments de droite à partir des données concernant la position

des maxima.

Cette méthode ne donne quasiment pas de fausses alarmes et le taux de détection est très élevé. Toutefois la transformée de Hough standard est dépendante d'un point de vue temps de calcul et espace mémoire, du nombre de points en entrée (points de contour de l'image initiale). D'autre part, le choix d'une résolution optimale, pour l'accumulateur peut être compliqué. Pour pallier ces problèmes, différentes méthodes ont été proposées ces dernières années.

La transformée de Hough rapide (FHT) et la transformée de Hough adaptative (ATH) :

L'idée est de faire une recherche itérative dans l'espace des paramètres avec une résolution de plus en plus fine. Dans la FHT, l'espace des paramètres est divisé récursivement en carrés ayant une résolution de plus en plus fine. Ainsi une transformée de Hough sera réalisée uniquement sur les carrés ayant une note supérieure à un certain seuil T. Ce processus est répété jusqu'à ce que la taille du carré atteigne sa valeur limite, ou jusqu'à ce que la note soit inférieure au seuil T.

L'avantage de cette méthode est de permettre une implémentation parallèle facile et d'avoir un certain degré d'adaptativité.

L'algorithme ATH adopte le même raisonnement que précédemment, dans le sens où il s'intéresse toujours à une recherche multirésolution. Mais, ici, on découpe l'image en imagette (par exemple 9*9). Et c'est précisément sur ces imagettes que l'on fait l'analyse multirésolution. Les maxima sont identifiés par seuillage et par l'analyse des imagettes connexes. Un inconvénient de cette méthode est l'extraction des droites par segment.

C.2.3 Les approches probabilistes

La complexité de la transformée de Hough est dépendante de la taille des données en entrée. Une façon de diminuer les temps de calcul et le stockage mémoire est de choisir des sous-ensembles de points de manière aléatoire.

La transformée de Hough aléatoire

On sélectionne de façon aléatoire des couples de points dans l'espace image. Si la distance entre ces points est suffisante (supérieure à un seuil), on continue sinon on fait une autre sélection. On détermine le couple de paramètres correspondant et on incrémente l'accumulateur. Des extensions de cette méthode existent :

The Random Hough Transform : RHT

La transformée de Hough aléatoire et dynamique : c'est une méthode itérative. Dans un premier temps une RHT est réalisée jusqu'à ce que le maximum global soit détecté dans l'espace des paramètres. On applique à nouveau une RHT, dans

un voisinage de la ligne détectée précédemment (sur un sous-ensemble de points de l'image initiale). Les points caractérisés par le maximum de l'accumulateur sont alors supprimés et on réitère le processus depuis le début.

La RHT à fenêtrage aléatoire utilise une information locale. On choisit de façon aléatoire un pixel dans l'image (point d'analyse), et une taille de fenêtre d'analyse, ainsi qu'un nombre maximum de sélections S. On sélectionne une paire de points dans cette fenêtre qui satisfasse la condition de distance. On détermine les paramètres dans l'espace dual et on incrémente l'accumulateur. Si la note attribuée à la cellule est inférieure au seuil, on recommence en choisissant deux points dans la fenêtre sachant que le nombre de sélection doit être inférieur à S, sinon on sélectionne à nouveau un autre point d'analyse. Si la note attribuée à la cellule est égale au seuil, on retient l'ensemble des paramètres possibles, et on retire les points sélectionnés de l'espace initial. Si on n'a pas atteint le nombre maximum de lignes à détecter, on choisit alors un autre point d'analyse et on itère le processus.

The Dynamic Combinatorial Hough Transform : DCHT

Une caractérisation ligne par ligne est utilisée. Pour cela, on choisit de façon aléatoire un point de référence dans l'image. Ainsi, avec tous les autres points de l'image, on peut déterminer à chaque fois un angle θ caractéristique. L'accumulateur sera dans ce cas un histogramme suivant θ. Dès que le seuil T est atteint on considère avoir détecté une droite. Celle-ci est alors effacée et on réitére le processus.

Probabilistic Hough Transform : ProbHT

Cette méthode utilise un sous-ensemble des points de contours choisis de manière aléatoire, dans l'image. Ces points sont alors utilisés comme entrée d'une transformée de Hough. Ainsi, le temps de calcul peut être considérablement réduit. Il est toutefois important de noter que les meilleures performances sont obtenues avec des sous-ensembles de petites tailles.

Cette technique diffère de la transformée de Hough aléatoire sur deux points. La transformation ne se fait ici que sur un sous-ensemble de points de contour. D'autre part, le principe d'accumulation est différent. On ne stocke pas dans l'accumulateur les valeurs pour chacun des points, mais on les regroupe par deux, ce qui se traduit par un stockage des droites.

Comparatifs

Une analyse comparative a été faite notamment sur une scène complexe réelle. La méthode la plus précise est sans aucun doute la SHT. Le problème est sa lenteur. Il est donc possible de faire un compromis sachant que souvent la rapidité se fera au détriment de la précision. Toutefois des techniques probabilistes, comme par exemple la ProbHT, donnent de bons résultats. Il en va de même pour la DCHT qui a en plus l'avantage de considérablement réduire la mémoire nécessaire.

La RWRHT est une méthode qui est également rapide, elle a l'avantage d'utiliser une information locale, et d'être adaptative, notamment par rapport à la taille de la fenêtre d'analyse.

C.3 Enveloppe convexe : algorithme de Graham

Cet algorithme utilisé pour déterminer l'enveloppe convexe d'un ensemble de points se décompose en trois phases :

- on se donne un point de référence P_{i0} appartenant à l'enveloppe convexe tel que le point le plus petit pour l'ordre lexicographique. On se donne également une direction D, par exemple, celle des x croissants.
- pour chaque point P_j du nuage, on considère l'angle entre D et $P_{i0}P_j$. Ces points sont triés par angles croissants.
- on réalise la marche de "Graham". Elle consiste à suivre les points dans l'ordre et à supprimer ceux qui forment un angle rentrant. Cette étape se fait en O(N).

C.4 Les mesures d'orientation

Ces directions sont calculées à partir des moments d'inertie centrés du deuxième ordre :

$$\mu_{xx} = \sum_{(i,j)\in\mathcal{R}} (j - j_G)^2$$
$$\mu_{yy} = \sum_{(i,j)\in\mathcal{R}} (i - i_G)^2$$
$$\mu_{xy} = \sum_{(i,j)\in\mathcal{R}} (i - i_G)(j - j_G)$$

i_G et j_G représentent les coordonnées du centre d'inertie de la région R. L'ensemble (i,j) décrit les coordonnées des pixels.

D'autre part, la direction α de l'axe d'inertie principal permet de caractériser l'orientation par rapport à l'horizontale.

$$\tan(2\alpha) = \frac{2\mu_{xy}}{\mu_{yy} - \mu xx}$$

Bien évidemment, cette mesure n'a de sens que si l'objet ne présente pas de symétrie de révolution, c'est-à-dire si $\mu_{yy} \neq \mu xx$.

C.5 Détecteur de Harris

Cet opérateur est basé sur la matrice M :

$$\begin{pmatrix} \widehat{I_x^2} & \widehat{I_x I_y} \\ \widehat{I_x I_y} & \widehat{I_y^2} \end{pmatrix}$$

où $\widehat{I}$ représente le résultat de la convolution de l'image I par une gaussienne. Les coins sont extraits par la recherche des maxima locaux de l'opérateur suivant :

$$C = det(M) - 0.04 Trace^2(M)$$

avec Trace= $(\widehat{I_x^2} + \widehat{I_y^2})^2$

L'attribut C permet de combiner l'information contour donnée par la trace de M et l'information angularité donnée par le déterminant de la matrice M.

C.6 Les ondelettes de Gabor

D. Gabor a proposé en 1946, une analyse de Fourier fenêtrée permettant une analyse temps-fréquence du signal. Le filtre utilisé n'est plus dans ce cas une sinusoïde (Fourier), mais une sinusoïde amortie.

Les ondelettes de Gabor ont la particularité d'avoir une fenêtre d'analyse de longueur fixe. L'enveloppe est constante et le nombre d'oscillations augmente avec la fréquence. La fenêtre optimale pour la localisation en temps est obtenue avec la fenêtre de Gauss. La fonction de base 1D s'écrit :

$$G_n(x) = e^{-\pi(\frac{-x^2}{a^2})} * e^{2i\pi\omega r cos(\theta-\phi)}$$

C'est la combinaison d'une enveloppe gaussienne de taille "a" et d'une exponentielle complexe de fréquence ω et d'orientation ϕ.

Le choix du triplet (a, ω, phi) définit une ondelette particulière. Dans la pratique [Casasent et Ye, 1997] pour la détection, le meilleur filtre obtenu est défini en 1D par :

$$g = \frac{3.32}{a}.e\frac{-\pi x^2}{d^2}.sin\frac{\pi x}{a}$$

C'est la partie imaginaire de la fonction de Gabor précédemment énoncée. Le coefficient $\frac{3.32}{a}$ correspond au coefficient de normalisation. le coefficient 2a définit la période du signal. Par conséquent la fréquence sera inversement proportionnelle à ce coefficient "a". Plus "a" sera grand, plus la fréquence sera faible et plus on analysera finement notre image. Inversement un coefficient "a" petit impose une grande fréquence, par conséquent une partie de l'information pourra être dissimulée.

Annexe D

Exemple : modèle du monde

Le modèle du monde obtenu pour une analyse directe dans l'image est présenté ci-dessous.

********** Modele du monde **********

```
    IMAGE : IMAGES/SOURCE/vol/vol.00.gif
Id= 0, taille = (640,480) * Concept :3
Nb Ech :8

    ******************************************************************
* Support :8
* estimation :74
* size :1
* zone :(Zone - 312 243 ;324 250
* nb pixel :0
* nb ech :1
Echantillon(s) :
ECH --- 3
Depart :POINT :312 :248
Haut :POINT :312 :243
Bas :POINT :324 :250
Nb Pix : 57
Ech Pere : 0
Pere : 3
Coefficient d'appartenance : 74.0005
Appartenance Gris : 0.703903
Appartenance Transition : 0.776107

Identificateur stat : 101 Moyenne : 121.632
Ecart : 29.6838
Variance : 881.129
Mini : 93
Maxi : 217

Identificateur stat : 102
Moyenne : 6.70175
Ecart : 1.13334
Variance : 1.28446
Mini : 4
Maxi : 8
    ******************************************************************
    ******************************************************************
* Support :7
* estimation :73
* size :1
* zone :(Zone - 276 252 ;281 257
* nb pixel :0
```

```
* nb ech :1
Echantillon(s) :
ECH --- 4
Depart :POINT :276 :256
Haut :POINT :276 :252
Bas :POINT :281 :257
Nb Pix : 26
Ech Pere : 0
Pere : 3
Coefficient d'appartenance : 73.3612
Appartenance Gris : 0.786293
Appartenance Transition : 0.680931

Identificateur stat : 101
Moyenne : 135.654
Ecart : 43.6431
Variance : 1904.72
Mini : 85
Maxi : 227

Identificateur stat : 102
Moyenne : 7.19231
Ecart : 0.938903
Variance : 0.881538
Mini : 5
Maxi : 8

    ********************************************************************
    ********************************************************************
* Support :6
* estimation :67
* size :1
* zone :(Zone - 364 241 ;367 247
* nb pixel :0
* nb ech :1
Echantillon(s) :
ECH --- 2
Depart :POINT :364 :244
Haut :POINT :364 :241
Bas :POINT :367 :247
Nb Pix : 23
Ech Pere : 0
Pere : 3
Coefficient d'appartenance : 67.8721
Appartenance Gris : 0.649133
Appartenance Transition : 0.708309

Identificateur stat : 101
Moyenne : 116.348
Ecart : 22.17
Variance : 491.51
Mini : 93
Maxi : 190

Identificateur stat : 102
Moyenne : 7.04348
Ecart : 1.06508
Variance : 1.13439
Mini : 5
Maxi : 8

    ********************************************************************
    ********************************************************************
* Support :5
* estimation :68
* size :1
* zone :(Zone - 471 206 ;478 210
* nb pixel :0
* nb ech :1
Echantillon(s) :
```

```
ECH --- 1
Depart :POINT :471 :208
Haut :POINT :471 :206
Bas :POINT :478 :210
Nb Pix : 32
Ech Pere : 0
Pere : 3
Coefficient d'appartenance : 68.2156
Appartenance Gris : 0.616744
Appartenance Transition : 0.747568

Identificateur stat : 101
Moyenne : 110.344
Ecart : 16.8381
Variance : 283.523
Mini : 85
Maxi : 147

Identificateur stat : 102
Moyenne : 6.84375
Ecart : 1.22104
Variance : 1.49093
Mini : 3
Maxi : 8

    ******************************************************************
    ******************************************************************
* Support :4
* estimation :62
* size :1
* zone :(Zone - 544 441 ;548 445
* nb pixel :0
* nb ech :1
Echantillon(s) :
ECH --- 8
Depart :POINT :544 :443
Haut :POINT :544 :441
Bas :POINT :548 :445
Nb Pix : 18
Ech Pere : 0
Pere : 3
Coefficient d'appartenance : 62.6446
Appartenance Gris : 0.622257
Appartenance Transition : 0.630636

Identificateur stat : 101
Moyenne : 113.667
Ecart : 2.42536
Variance : 5.88235
Mini : 112
Maxi : 117

Identificateur stat : 102
Moyenne : 3.55556
Ecart : 1.75641
Variance : 3.08497
Mini : 2
Maxi : 7

    ******************************************************************
    ******************************************************************
* Support :3
* estimation :65
* size :1
* zone :(Zone - 467 468 ;471 471
* nb pixel :0
* nb ech :1
Echantillon(s) :
ECH --- 9
Depart :POINT :467 :469
```

```
Haut :POINT :467 :468
Bas :POINT :471 :471
Nb Pix : 15
Ech Pere : 0
Pere : 3
Coefficient d'appartenance : 65.8192
Appartenance Gris : 0.615264
Appartenance Transition : 0.701121

Identificateur stat : 101
Moyenne : 113.667
Ecart : 2.4399
Variance : 5.95312
Mini : 112
Maxi : 117

Identificateur stat : 102
Moyenne : 3.93333
Ecart : 1.90738
Variance : 3.6381
Mini : 2
Maxi : 7

    ********************************************************************
    ********************************************************************
* Support :2
* estimation :63
* size :1
* zone :(Zone - 565 438 ;574 443
* nb pixel :0
* nb ech :1
Echantillon(s) :
ECH --- 7
Depart :POINT :565 :441
Haut :POINT :565 :438
Bas :POINT :574 :443
Nb Pix : 35
Ech Pere : 0
Pere : 3
Coefficient d'appartenance : 63.2505
Appartenance Gris : 0.642856
Appartenance Transition : 0.622154

Identificateur stat : 101
Moyenne : 113.857
Ecart : 2.45137
Variance : 6.00919
Mini : 112
Maxi : 117

Identificateur stat : 102
Moyenne : 3.48571
Ecart : 1.4627
Variance : 2.1395
Mini : 2
Maxi : 7

    ********************************************************************

    ********************************************************************
* Support :1
* estimation :67
* size :1
* zone :(Zone - 578 434 ;586 441
* nb pixel :0
* nb ech :1
Echantillon(s) :
ECH --- 6
Depart :POINT :578 :439
Haut :POINT :578 :434
```

```
Bas :POINT :586 :441
Nb Pix : 33
Ech Pere : 0
Pere : 3
Coefficient d'appartenance : 67.1246
Appartenance Gris : 0.641203
Appartenance Transition : 0.701289

Identificateur stat : 101
Moyenne : 113.818
Ecart : 2.4425
Variance : 5.96582
Mini : 112
Maxi : 117

Identificateur stat : 102
Moyenne : 3.90909
Ecart : 2.08212
Variance : 4.33523
Mini : 2
Maxi : 7

    ******************************************************************

    ******************************************************************
* Support :0
* estimation :63
* size :1
* zone :(Zone - 603 431 ;611 437
* nb pixel :0
* nb ech :1
Echantillon(s) :
ECH --- 5
Depart :POINT :603 :435
Haut :POINT :603 :431
Bas :POINT :611 :437
Nb Pix : 34
Ech Pere : 0
Pere : 3
Coefficient d'appartenance : 63.0845
Appartenance Gris : 0.64255
Appartenance Transition : 0.61914

Identificateur stat : 101
Moyenne : 113.912
Ecart : 2.46625
Variance : 6.08239
Mini : 112
Maxi : 117

Identificateur stat : 102
Moyenne : 3.47059
Ecart : 1.44044
Variance : 2.07487
Mini : 2
Maxi : 7

    ******************************************************************
```

Bibliographie

[Acheroy et Mees, 1991] M. Acheroy et W. Mees. Target identification by means of adaptative neural networks in thermal infrared images. *Stochastic and Neural Methods in Signal Processing, Image Processing, and Computer Vision*, 1569 :121–132, 1991.

[Auborn *et al.*, 1993] J.E. Auborn, J.M. Fuller, et H.M. McCauley. Target detection by co-occurence matrix segmentation and its hardware implementation. *Optical Engineering*, 32(11) :2698–2702, Novembre 1993.

[Bajcsy, 1988] R. Bajcsy. Active perception. *Proceedings of the IEEE*, 76 :996–1005, 1988.

[Barrow *et al.*, 1977] H.G. Barrow, J.M. Tenenbaum, R.C. Bolles, et H.C. Wolf. Parametric correspondence and chamfer matching : two new techniques for image matching. *IJCAI77*, pages 659–663, 1977.

[Becker et Naïm, 1999] A. Becker et P. Naïm. *Les réseaux bayésiens, modèles graphiques de connaissance*. Eyrolles, eyrolles édition, Juin 1999.

[Beveridge *et al.*, 1997] J.R. Beveridge, B.A. Draper, M.R. Stevens, A. Hanson, et K. Siejko. A coregistration approach to multisensor target recognition with extensions to exploit digital elevation map data. Rapport Technique CS-97-107, Colorado State University, Septembre 1997.

[Bhanu et Jones, 1992] B. Bhanu et T.L. Jones. Image understanding research for automatic target recognition. *IUW*, pages 249–254., Janvier 1992.

[Bhanu, 1986] B. Bhanu. Automatic Target Recognition : State of the Art Survey. *IEEE transactions on aerospace and electronic systems*, 22(4) :364–379, Juillet 1986.

[Block et Maitre, 1994] I. Block et H. Maitre. Fusion de données en traitement d'images : modèles d'information et décisions. *Traitement du signal*, 11(6) :435–446, 1994.

[Borgefors, 1988] G. Borgefors. Hierarchical chamfer matching : a parametric edge matching algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-10(6) :849–865, 1988.

[Boucher, 1999] A. Boucher. *Une approche décentralisée et adaptative de la gestion d'information en vision*. Thèse de Doctorat, Université Joseph Fourier, Janvier 1999.

[Bouillot, 1999] A. Bouillot. Développement d'une architecture multi-agent pour le contrôle d'un robot mobile. CTA 99 R 013, DGA/DCE/CTA/GIP, 1999.

[Brémond et Thonnat, 1996] F. Brémond et M. Thonnat. Interprétation de séquences d'images et incertitude. *Rencontres sur la logique floue et ses applications (LFA)*, Nancy, Décembre 1996.

[Buede et Girardi, 1997] D.M. Buede et P. Girardi. A target identification comparison of bayesian and dempster-shafer multisensor fusion. *IEEE transactions on systems man and cybernetics - part A : systems and humans*, 27(5) :569–577, Septembre 1997.

[Burger et Bhanu, 94] W. Burger et B. Bhanu. Signal-to-symbol conversion for structural object recognition using hidden markov models. *Darpa Image Understanding Workshop*, 94.

[Buxton et Gong, 1995] H. Buxton et S. Gong. Advanced visual surveillance using bayesian networks. *Artificial Intelligence*, 78(1-2) :431–459, Juin 1995.

[Casasent et Ye, 1997] D. Casasent et A. Ye. Detection filters and algorithm fusion for ATR. *IEEE Transactions on Image Processing*, 6(1) :114–125, Janvier 1997.

[Castello *et al.*, 2000] G. Castello, J. Boyce, et M. Sandler. Regularized CDWT optical flow applied to moving target detection in IR imagery. *Machine Vision and Applications*, 11(6) :277–288, 2000.

[Crevier et Lepage, 1997] D. Crevier et R. Lepage. Knowledge-based image understanding systems : a survey. *Computer Vision and Image Understanding*, 67(2) :161–185, 1997.

[Dalgalarrondo, 2001] A. Dalgalarrondo. *Intégration de la fonction perception dans une architecture de contrôle de robot mobile autonome.* Thèse de Doctorat, Université de Paris-Sud, Orsay, Janvier 2001.

[Das et Bhanu, 1998] S. Das et B. Bhanu. A system for model-based object recognition in perspective aerial images. *Pattern Recognition*, 31(4) :465–491, 1998.

[Dattatreya et Kanal, 1985] G.R. Dattatreya et L.N. Kanal. *Progress in Pattern Recognition*, chapitre Decision Tree in Pattern Recognition, pages 189–239. L. Kanal and A. Rosenfeld, New-York, Elsevier Science édition, 1985.

[Dearden *et al.*, 1998] R. Dearden, N. Friedman, et S. Russell. Bayesian Q-learning. *fifteenth National Conference on Artificial Intelligence*, pages 761–768, 1998.

[Der et Challappa, 1997] S. Z. Der et R. Challappa. Probe based automatic target recognition in infrared imagery. *IEEE Transactions on Image Processing*, 6(1) :92–102, Janvier 1997.

[Dickinson *et al.*, 1992] S. J. Dickinson, A. P. Pentland, et A. Rosenfeld. From Volumes to Views : an Approach to 3-D Object Recognition. *Computer Vision, Graphics and Image process : Image Uderstanding*, 55(2) :130–154, Mars 1992.

[Dillion et Caelli, 1998] C. Dillion et T. Caelli. Learning image annotation : the CITE system. *Videre : Journal of Computer Vision Research*, 1(2) :90–121, Août 1998.

[Djian, 1997] D. Djian. *Contribution à l'analyse de scènes par vision active : utilisation de réseaux bayésiens.* Thèse de Doctorat, Ecole des Mines de Paris, Février 1997.

[Drake et Kim, 1995] K.C. Drake et R.Y. Kim. Hierarchical integration of sensor data and contextual information for automatic target recognition. *Applied Intelligence*, 5 :269–290, 1995.

[Draper *et al.*, 1989] B. Draper, R. Collins, J. Brolio, A. Hanson, et E. Riseman. The SCHEMA system. *International Journal of Computer Vision*, 2 :209–250, 1989.

[Draper *et al.*, 1996] B.A. Draper, A.R. Hanson, et E.M. Riseman. Knowledge-directed vision : control, learning and integration. *Proceedings of IEEE*, 84(11) :1625–1681, 1996.

[Duda *et al.*, 2001] R.O. Duda, P.E. Hart, et D.G. Stork. *Pattern classification.* John Wiley Intersciences, 2001.

[Dudgeon et Lacoss, 1993] D.E. Dudgeon et R.T. Lacoss. An overview of automatic target recognition. *The lincoln laboratory journal*, 6(1) :3–10, 1993.

[Ealet *et al.*, 1999] F. Ealet, E. Touboul, B. Collin, G. Sella, et C. Garbay. Un système multi-agent pour la détection, la reconnaissance et l'identification de cibles. *ORASIS99*, pages 289–298, Aussois, France, Avril 1999.

[Ealet *et al.*, 2000] F. Ealet, B. Collin, G. Sella G, et C. Garbay. Multi-agent architecture for scene interpretation. *SPIE, Enhanced and synthetic vision 2000*, Orlando, Floride, Avril 2000.

[Ernisse *et al.*, 1997] B. Ernisse, S.K. Rogers, M.P. DeSimio, et R.A. Raines. Complete automatic target cuer/recognition system for tactical forward-looking infrared images. *Optical Engineering*, 36(9) :2593–2603, Septembre 1997.

[Espinal *et al.*, 1998] F. Espinal, B.D. Jawerth, et T. Kubota. Wavelet-based fractal signature analysis for automatic target recognition. *Optical Engineering*, 37(1) :166–174, Janvier 1998.

[Fabiani, 1996] P. Fabiani. *Représentation dynamique de l'incertain et stratégie de perception pour un système autonome en environnement évolutif.* Thèse de Doctorat, Ecole national supérieure de l'aéronautique et de l'espace, 1996.

[Ferber, 1995] J. Ferber. *Les systèmes multi-agents : vers une intelligence collective.* InterEdition, Paris, France, 1995.

[Garbay, 2000] C. Garbay. Architecture logicielle et contrôle dans les systèmes de vision. *Les systèmes de vision*, éditeur J.M. Jolion. Hermes, 2000.

[Gatepaille et Brunessaux, 1996] S. Gatepaille et S. Brunessaux. Dossier I.A. et fusion de données. *Bulletin de l'AFIA*, 24 :21–30, 1996.

[Gupta *et al.*, 1995] L. Gupta, T. Sortrakul, A. Charles, et P. Kisatsky. Robust automatic target recognition using a localized boundary representation. *Pattern Recognition*, 28(10) :1587–1598, 1995.

[Hamilton et Kipp, 1993] M. Hamilton et T. Kipp. Model-based multisensor fusion, Novembre 1993.

[Hecht-Nielsen *et al.*, 1994] R. Hecht-Nielsen, M. Shao, et S. Wang. Fast k-nn search for robust atr object matching. *Image Understanding Workshop*, 94 :889–894, 1994.

[Henocq et Burel, 1993] H. Henocq et G. Burel. Reconnaissance d'objets 3D par analyse de fourier de la silhouette. *Revue Technique Thomson-CSF*, 25(1) :107–125, Mars 1993.

[Hoc, 1996] J. Hoc. *Supervision et contrôle de processus : la cognition en situation dynamique*. PUG, Grenoble, 1996.

[Howard et Matheson, 1984] R. Howard et J. Matheson. *READINGS on The Principles and Applications of Decision Analysis*, chapitre Influence diagrams. Strategic Decisions Group, 1984.

[Hsu *et al.*, 1994] S. Hsu, J. Pearson, P. Sajda, et C. Spence. Neural network/pyramid architectures that learn target context. *ARPA94*, volume II, pages 853–862, 1994.

[Huang et Russell, 1998] T. Huang et S. Russell. Object identification : a bayesian analysis with application to traffic surveillance. *Artificial Intelligence*, 103 :77–93, 1998.

[Jain *et al.*, 1997] A. K. Jain, N. K. Ratha, et S. Lakshmanan. Object detection using Gabor filter. *Pattern Recognition*, 30(2) :295–309, 1997.

[Jensen *et al.*, 1990] F.V. Jensen, S.L. Lauritzen, et K.G. Olesen. Bayesian updating in causal probabilistic networks by local computation. *Computational Statistics Quarterly*, 4 :269–282, 1990.

[Jonquières, 2000] S. Jonquières. *Application des réseaux Bayesiens à la reconnaissance active d'objets 3D : contribution à la saisie d'objets.* Thèse de Doctorat, Institut National Polytechnique de Toulouse, Janvier 2000.

[Kaelbling *et al.*, 1996] L. Kaelbling, M. Littman, et A. Moore. Reinforcement learning : a survey. *Journal of Artificial Intelligence Research*, 4 :237–285, 1996.

[Keerthi et Ravindran, 1994] S. Sathiya Keerthi et B. Ravindran. A tutorial survey of reinforcement learning. *Sadhana*, 19(6) :851–889, Décembre 1994.

[Kornprobst *et al.*, 1997] . Kornprobst, R. Deriche, et G. Aubert. Image sequence restoration : a PDE based coupled method for image restoration and motion segmentation. Rapport Technique 3308, INRIA, Novembre 1997.

[Lamdan et Wolfson, 1988] Y. Lamdan et H. J. Wolfson. Geometric hashing : a general and efficient model-based recognition scheme. *2nd IEEE International Conference on Computer Vision (ICCV'88)*, pages 238–249, Tampa, Floride, U.S.A., Décembre 1988.

[Lefevre, 1994] V. Lefevre. *Système multi-agent d'aide à la photo-interprétation sur architecture multi-processeur.* Thèse de Doctorat, Université de Paris XI, Novembre 1994.

[Lefèvre *et al.*, 1996] V. Lefèvre, Y. Collet, S. Philipp, et S. Brunessaux. Un système multi-agent pour la fusion de données en analyse d'images. *Traitement du signal*, 13(1) :99–111, 1996.

[Leung et Huang, 1991] M.K. Leung et T.S. Huang. An integrated approach to 3D motion analysis and object recognition. *Pattern Analysis and Machine Intelligence*, 13(10) :1075–1084, Octobre 1991.

[Li *et al.*, 1998] B. Li, Q. Zheng, S. Der, R. Chellappa, N.M. Nasrabadi, L.A. Chan, et L.C. Wang. Empirical evaluation of neural, statistical and model-based approaches to flir atr. *UMD*, 1998.

[Marchand et Chaumette, 1998] E. Marchand et F. Chaumette. Stratégies de perception par vision active : une approche par réseaux bayesiens. *Revue d'Intelligence Artificielle*, 12(3) :345–376, Septembre 1998.

[Mees, 2000] W. Mees. *Contribution à l'analyse de scènes ; Application aux images satellitaires multi-spectrales, haute résolution.* Thèse de Doctorat, Université Henri Poincaré, Nancy 1, Janvier 2000.

[Moissinac *et al.*, 1995] H. Moissinac, H. Maitre, et I. Block. Markov random fields and graphs for uncertainty management and symbolic data fusion in a urban scene interpretation. *SPIE Image and Signal Processing for Remote Sensing*, 2579 :298–309, 1995.

[Moissinac-Massenat, 1996] Henri Moissinac-Massenat. *Utilisation conjointe d'informations symboliques et de mesures numériques dans la prise de décision en traitement d'images.* Thèse de Doctorat, Ecole nationale supérieure des télécommunications, Paris, Mars 1996.

[Murthy, 1998] S. K. Murthy. Automatic construction of decision trees from data : A multi-disciplinary survey. *Data Mining and Knowledge Discovery*, 2(4) :345–389, 1998.

[Nair et Aggarwal, 2000] D. Nair et J.K. Aggarwal. Bayesian recognition of targets by parts in second generation forward looking infrared images. *Image and Vision Computing*, 18 :849–864, 2000.

[Nandhakumar *et al.*, 1997] N. Nandhakumar, J.D. Michel, D.G. Arnold, G.A. Tsihrintzis, et V. Velten. Robust thermophysics-based interpretation of radiometrically uncalibred IR images for ATR and site change detection. *IEEE Transactions on Image Processing*, 6(1) :65–78, Janvier 1997.

[Nifle, 1998] A. Nifle. *Modélisation comportementale en fusion de données ; Application à l'identification d'objets ou de situations.* Thèse de Doctorat, Université de Orsay, 1998.

[Nigro *et al.*, 2000] J.M. Nigro, S. Loriette-Rougegrez, M. Rombaut, et I. Jarkass. Driving situation recognition in the CASSICE project towards an uncertain management. *ITSC*, 2000.

[Olson et Huttenlocher, 1997] C.F. Olson et D.P. Huttenlocher. Automatic target recognition by matching oriented edge pixel. *IEEE Transactions on Image Processing*, 6(1) :103–113, Janvier 1997.

[Pampagnin, 1990] L.-H. Pampagnin. *Reconnaissance d'objets tridimensionnels en perception monoculaire et multisensorielle - Application à la robotique spatiale.* Thèse de Doctorat, LAAS-CNRS, Toulouse, Octobre 1990.

[Pearl, 1988] J. Pearl. *Probabilistic reasoning in intelligent systems : Network of plausible inference*. Morgan Kauffman, 1988.

[Preparara et Shamos, 1985] F.P. Preparara et M.I. Shamos. *Computational Geometry, an introduction*. Springer-Verlag, 1985.

[Prokop et Reeves, 1992] R.J. Prokop et A.P. Reeves. A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP : Graphical, Models and Image Processing*, 54(5) :438–460, Septembre 1992.

[Rao et Ballard, 1995] R.P. Rao et D.H. Ballard. An active vision architecture based on iconic representations. *Artificial Intelligence Journal*, 78(1-2) :461–505, 1995.

[Rao, 1993] K. Rao. Extracting salient contours for target recognition : algorithm and performance evaluation. *Optical Engineering*, 32(11) :2690–2697, Novembre 1993.

[Ratches *et al.*, 1997] J.A. Ratches, C.P. Walters, R. G. Buser, et B.D. Guenther. Aided and automatic target recognition based upon sensory inputs from image forming systems. *IEEE transactions on pattern analysis and machine intelligence*, 19(4) :1004–1019, Septembre 1997.

[Rigoutsos et Hummel, 1995] I. Rigoutsos et R. Hummel. A bayesian approach to model matching with geometric hashing. *Computer vision and image understanding : CVIU*, 62(1) :11–26, 1995.

[Rimey, 1993] R. D. Rimey. Control of selective perception using bayes nets and decision theory. Rapport technique, The university of Rochester, Computer science department, Rochester, New York 14627, Décembre 1993.

[Roberts et Brown, 1994] B. Roberts et C. Brown. Adaptative and control in an ATR system. *IUW*, pages 467–479, 1994.

[Rogers *et al.*, 1995] S.K. Rogers, J.M. Colombi, C.E. Martin, K.H. Fielding J.C. Gainey, T.J. Burns, D.W. Ruck, M. Kabrisky, et M. Oxley. Neural networks for automatic target recognition. *Neural network*, 8(7-8) :1153–1184, 1995.

[Rombaut *et al.*, 2000] M. Rombaut, S. Loriette-Rougegrez, J.M. Nigro, et I. Jarkass. Temporal sequence recognition using incertain data. *FUSION'2000*, Paris, 2000.

[Ropert, a paraitre] V. Ropert. *Vers un système de vision basé sur une étude comportementale d'algorithmes*. Thèse de Doctorat, Université Paris VII, à paraître.

[Roth, 1990] M. W. Roth. Survey of neural network technology for automatic target recognition. *IEEE Transactions on Neural Networks*, 1 :28–43, 1990.

[Safavin et Landgrebe, 1991] S.R. Safavin et D. Landgrebe. A Survey of Decision Tree Classifier Methodology. *IEEE Trans. System, Man and Cybernetics*, 21 :660–674, Mai 1991.

[Sandakly, 1995] F. Sandakly. *Contribution à la mise en œuvre d'une architecture à base de connaissances pour l'interprétation de scènes 2D et 3D*. Thèse de Doctorat, Université de Nice-Sophia Antipolis, Juillet 1995.

[Schmid et Mohr, 1996] C. Schmid et R. Mohr. Mise en correspondance par invariants locaux. *Traitement du signal*, 13(6) :591–606, 1996.

[Scribner *et al.*, 2000] D. Scribner, P. Warren, et J. Schuler. Extending color vision methods to bands beyond the visible. *Machine Vision and Applications*, 11(6) :306–312, 2000.

[Serfaty-Dutron et Blanc-Talon, 1999] V. Serfaty-Dutron et J. Blanc-Talon. Panorama et évolution des sytèmes d'interprétation de scène. Rapport Technique 46, DCE/CTA/GIP, Arcueil, 1999.

[Spiegelhalter et Lauritzen, 1990] D.S. Spiegelhalter et S.L. Lauritzen. Sequential updating of conditional probabilities on directed graphical structures. *Networks*, 20 :579–605, 1990.

[Stassopoulou et Caelli, 2000] A. Stassopoulou et T. Caelli. Building detection using bayesian networks. *International journal of pattern recognition and artificial intelligence*, 14(6) :715–733, 2000.

[Stevens et Beveridge, 1997] M.R. Stevens et J.R. Beveridge. Precise matching of 3D target models to multisensor data. *IEEE Transactions on Image Processing*, 6(1) :126–142, Janvier 1997.

[Stevens, 1999] M.R. Stevens. *Reasoning about object apparence in the context of a scene.* Thèse de Doctorat, Colorado State University, 1999.

[Suryadi et Gmytrasiewicz, 1999] D. Suryadi et P. Gmytrasiewicz. Learning models of other agents using influence diagram. *Proceedings of the 1999 International Conference on User Modeling*, pages 223–232, 1999.

[Touboul, 1998] E. Touboul. *Modèle d'architecture pour la mise en œuvre d'un système de vision générique.* These de doctorat, Conservatoire national des arts et metiers, 1998.

[Tsai, 1996] F. C. D. Tsai. A probabilistic approach to geometric hashing using line features. *Computer vision and image understanding : CVIU*, 63(1) :182–195, 1996.

[Verly et Delanoy, 1996] J. G. Verly et R. L. Delanoy. Model-Based Automatic Target Recognition (ATR) System for Forwardloking Groundbased and Airbone Imaging Laser Radar (LADAR). *Proceedings of the IEEE*, 84(2) :126–163, Février 1996.

[Wang *et al.*, 1998] L. Wang, S.Z. Der, et N.M. Nasrabadi. Composite classifiers for automatic target recognition. *Optical Engineering*, 37(3) :858–868, Mars 1998.

[Wu et Bhanu, 1997] X. Wu et B. Bhanu. Gabor wavelet representation for 3D object recognition. *IEEE Transactions on Image Processing*, 6(1) :47–64, Janvier 1997.

[Young *et al.*, 1998] S. S. Young, P. D.Scott, et C. Bandera. Foveal automatic target recognition using a multiresolution neural network. *IEEE Transactions on Image Processing*, 7(8) :1122–1135, Août 1998.

[Zhang, 1993] Z. Zhang. Le problème de la mise en correspondance : l'état de l'art. Rapport Technique 2146, INRIA, Décembre 1993.