

Université des Sciences et Technologies de Lille

Perception de l'environnement par stéréovision  
Application à la sécurité dans les systèmes de transports  
terrestres

Mémoire présenté  
en vue de l'obtention de

**l'Habilitation à Diriger des Recherches**

par  
Yassine Ruichek  
Maître de conférences à l'UTBM

Soutenu publiquement le 29 septembre 2005, devant le jury composé de :

Olivier Colot	Professeur à l'USTL	Président
Chuck Thorpe	Professeur à Carnegie Mellon University	Rapporteur
Bernard Dubuisson	Professeur à l'UTC	Rapporteur
Roland Chapuis	Professeur à l'Université Blaise-Pascal	Rapporteur
Djemel Ziou	Professeur à l'Université de Sherbrooke	Examineur
Abderrafiaa Koukam	Professeur à l'UTBM	Examineur
Jack-Gérard Postaire	Professeur à l'USTL	Examineur

## Remerciements

J'ai été constamment soutenu par l'amitié et la confiance de Monsieur J.-G Postaire avec qui j'ai effectué mes premiers pas vers la carrière d'enseignant-chercheur. Nos collaborations n'ont jamais arrêté de se développer tout au long de mon parcours universitaire. Je le remercie beaucoup pour le rôle clé qu'il a joué dans ce parcours.

Monsieur A. Koukam m'a accueilli au sein de son équipe en tant que Maître de Conférences et m'a fortement soutenu pour développer une nouvelle thématique au sein du laboratoire Systèmes et Transports. Il m'a donné les moyens de travailler dans d'excellentes conditions, je lui en suis très reconnaissant.

Messieurs C. Thorpe, B. Dubuisson et R. Chapuis ont accepté spontanément de juger mon travail. Je les remercie beaucoup du temps qu'ils ont pris pour la lecture et l'analyse de ce mémoire.

Je suis très honoré par la présence de Monsieur D. Ziou qui a survolé l'océan atlantique pour participer à ce jury. Je lui en suis très reconnaissant.

Je remercie chaleureusement Monsieur O. Collot qui a bien voulu faire partie du jury et assumer le rôle de président.

Je remercie tous les chercheurs (H. Issa, D. Dooze, M. Hariti, T. Yahiaoui) et collègues (C. Vieren, P. Vannoorenberghe, J.-C. Burie, L. Macaire, F. Cabestaing, L. Khoudour, S. Ambellouis, F. Gechter) avec qui j'ai eu le plaisir de travailler et partager des moments agréables.

P. Kmiotek et S. Nogueira, deux nouveaux doctorants, ont accepté de partager leurs expériences de recherche avec moi. Je les remercie pour leur confiance.

Je remercie A. Pignon, secrétaire du LAGIS, A. Chevalier, M. Perrin, secrétaires du laboratoire SeT, et N. Zaidi, secrétaire du département GI, pour leur aide quotidienne, toujours dans la bonne humeur.

Je remercie V. Hilaire, M. Bakhouya et A. Berdai pour l'aide qu'ils m'ont apportée pour apprendre à utiliser l'éditeur de texte LaTeX.

Mes remerciements vont également à tous mes collègues et amis pour leur soutien.

---

# TABLE DES MATIÈRES

<b>Introduction</b>	<b>1</b>
<b>Notice individuelle</b>	<b>14</b>
<b>Activités d’enseignement</b>	<b>18</b>
<b>Activités de recherche</b>	<b>28</b>

<b>Chapitre 1</b> <b>Perception par vision stéréoscopique</b>
--

1.1	Introduction . . . . .	39
1.2	Techniques pour la perception de l’environnement . . . . .	40
1.2.1	Approche active . . . . .	41
1.2.2	Approche passive . . . . .	42
1.3	Processus de stéréovision . . . . .	42
1.3.1	Décomposition du processus de stéréovision . . . . .	43
1.3.2	Principes généraux . . . . .	43
1.3.3	Aspects fondamentaux de l’appariement . . . . .	44
1.3.3.1	Fonction d’appariement . . . . .	44
1.3.3.2	Complexité du problème de l’appariement . . . . .	45
1.3.3.3	Dilemme précision – appariement . . . . .	45
1.4	Aspects géométriques de la stéréovision . . . . .	46

---

1.4.1	Géométrie du capteur stéréoscopique . . . . .	46
1.4.2	Calibrage . . . . .	46
1.4.3	Droites épipolaires . . . . .	47
1.4.4	Configuration idéale . . . . .	48
1.4.5	Reconstruction tridimensionnelle . . . . .	49
1.5	Les primitives . . . . .	50
1.5.1	Les propriétés requises . . . . .	50
1.5.2	Types de primitives . . . . .	51
1.5.2.1	Les points d'intérêt . . . . .	51
1.5.2.2	Les contours . . . . .	52
1.5.2.3	Les régions . . . . .	53
1.5.2.4	Les contours actifs . . . . .	54
1.6	Les critères de ressemblance . . . . .	54
1.7	Les contraintes . . . . .	56
1.7.1	Les contraintes locales . . . . .	56
1.7.1.1	La contrainte épipolaire . . . . .	56
1.7.1.2	La contrainte d'orientation . . . . .	56
1.7.1.3	La contrainte de disparité maximale et/ou minimale . . . . .	57
1.7.2	Les contraintes globales . . . . .	57
1.7.2.1	La contrainte d'unicité . . . . .	57
1.7.2.2	La contrainte d'ordre . . . . .	58
1.7.2.3	La contrainte de continuité de la disparité . . . . .	61
1.7.2.4	La contrainte de continuité figurale . . . . .	62
1.8	Mise en correspondance . . . . .	62
1.8.1	Méthodes surfaciques . . . . .	63
1.8.2	Méthodes à base de primitives . . . . .	63
1.8.3	Techniques de mise en correspondance . . . . .	64
1.8.3.1	Techniques corrélatives . . . . .	64
1.8.3.2	Techniques hiérarchiques . . . . .	65
1.8.3.3	Techniques basées sur la programmation dynamique . . . . .	66
1.8.3.4	Techniques d'étiquetage par relaxation . . . . .	67
1.8.3.5	Techniques par isomorphisme de graphes . . . . .	69
1.8.3.6	Techniques par prédiction et vérification d'hypothèses . . . . .	71
1.8.3.7	Techniques par réseaux de neurones . . . . .	72

1.8.3.8	Techniques par algorithmes génétiques . . . . .	74
1.8.3.9	Techniques utilisant le recuit simulé . . . . .	75
1.8.3.10	Autres techniques . . . . .	75
1.9	Conclusion . . . . .	76

## Chapitre 2

### Mise en correspondance par réseau de neurones de Hopfield

2.1	Introduction . . . . .	79
2.2	Stéréovision linéaire . . . . .	80
2.3	Extraction des primitives . . . . .	82
2.4	Mise en correspondance . . . . .	86
2.4.1	Codage du problème . . . . .	87
2.4.2	Contraintes de mise en correspondance . . . . .	88
2.4.2.1	Contraintes locales . . . . .	88
2.4.2.2	Contraintes globales . . . . .	90
2.4.3	Fonction objective . . . . .	91
2.4.3.1	Fonction objective représentant la contrainte d'unicité . . . . .	92
2.4.3.2	Fonction objective représentant la contrainte d'ordre . . . . .	93
2.4.3.3	Fonction objective représentant la contrainte de continuité de la disparité . . . . .	94
2.4.3.4	Fonction objective totale . . . . .	96
2.4.4	Procédure neuronale pour la minimisation de la fonction objective . . . . .	96
2.4.4.1	Construction du réseau . . . . .	99
2.4.4.2	Processus de minimisation . . . . .	102
2.4.5	Analyse de l'état final du réseau . . . . .	104
2.4.6	Choix des paramètres . . . . .	105
2.4.6.1	Choix des paramètres $\theta$ et $\alpha$ . . . . .	105
2.4.6.2	Choix des coefficients de pondération $K_u$ , $K_o$ et $K_c$ . . . . .	106
2.4.6.3	Choix des paramètres $\tau$ et $\lambda$ . . . . .	106
2.5	Résultats de mise en correspondance . . . . .	107
2.5.1	Séquence 1 : Piéton évoluant devant le stéréoscope . . . . .	109
2.5.2	Séquence 2 : Approche d'un véhicule . . . . .	112
2.6	Mise en correspondance par séparation des points de contour . . . . .	116
2.7	Résultats au sein de l'équipe "Image et Décision" . . . . .	117
2.8	Conclusion . . . . .	119

## Chapitre 3

### Mise en correspondance par algorithmes génétiques

3.1	Introduction . . . . .	122
3.2	Codage binaire . . . . .	122
3.3	Evaluation des chromosomes binaires . . . . .	124
3.4	Mise en oeuvre de l'algorithme génétique . . . . .	126
3.5	Résultats de mise en correspondance issus de l'AGB . . . . .	129
3.5.1	Choix des paramètres . . . . .	129
3.5.1.1	Paramètres intervenant dans la fonction d'évaluation . . .	129
3.5.1.2	Paramètres intervenant dans l'algorithme génétique . . . .	130
3.5.2	Résultats de mise en correspondance . . . . .	131
3.5.3	Conclusion . . . . .	132
3.6	Limitations du codage binaire . . . . .	133
3.7	Nouveau codage : le codage entier . . . . .	135
3.7.1	Construction du nouveau codage . . . . .	135
3.7.2	Intérêt du nouveau codage . . . . .	138
3.8	Evaluation des chromosomes entiers . . . . .	143
3.9	Réglage de la taille de la population et du nombre de générations . . . . .	145
3.10	Résultats de mise en correspondance obtenus par l'AGE . . . . .	146
3.11	Conclusion . . . . .	148

## Chapitre 4

### Mise en correspondance rapide

4.1	Introduction . . . . .	150
4.2	Stratégie de recherche hiérarchique . . . . .	150
4.2.1	Principe de base . . . . .	152
4.2.2	Définitions et formalisme . . . . .	154
4.2.3	Sélection des points de contour . . . . .	155
4.2.4	Module d'appariement élémentaire . . . . .	157
4.2.5	Conditions d'arrêt de la recherche hiérarchique . . . . .	158
4.2.5.1	Condition d'arrêt 1 . . . . .	158
4.2.5.2	Condition d'arrêt 2 . . . . .	158
4.2.6	Implantation recursive de la stratégie hiérarchique . . . . .	159
4.2.6.1	Construction de l'arbre de recherche . . . . .	160

---

4.2.6.2	Extraction des couples appariés . . . . .	161
4.2.7	Evaluation de la stratégie de recherche hiérarchique . . . . .	162
4.2.7.1	Séquence "Piéton" . . . . .	162
4.2.7.2	Séquence "Approche d'un véhicule" . . . . .	166
4.2.8	Conclusion sur la recherche hiérarchique . . . . .	166
4.3	Stratégie de vote . . . . .	168
4.3.1	Etape 1 : Détermination des couples électeurs d'un couple candidat	170
4.3.2	Etape 2 : Calcul du score d'un couple candidat . . . . .	171
4.3.3	Extraction des appariements . . . . .	172
4.3.4	Evaluation de la stratégie de vote . . . . .	172
4.3.5	Conclusion sur la méthode de vote . . . . .	175
4.4	Conclusion . . . . .	176
<b>Conclusion et perspectives</b>		<b>178</b>
<b>Annexe 1</b>		
	<b>Activités d'enseignement à l'UTBM</b>	<b>183</b>
<b>Annexe 2</b>		
	<b>Activités d'encadrement</b>	<b>188</b>
<b>Annexe 3</b>		
	<b>Liste des publications</b>	<b>192</b>
<b>Bibliographie</b>		<b>199</b>
<b>Quelques publications</b>		<b>217</b>

---

# TABLE DES FIGURES

1	Répartition des enseignements par matière. . . . .	22
2	Répartition des enseignements par Cours, TD et TP. . . . .	22
3	Répartition des enseignements par niveau. . . . .	23
4	Répartition par type d'enseignement. . . . .	23
1.1	Configuration générale d'un capteur stéréoscopique. . . . .	47
1.2	Configuration particulière du système de prise de vue. . . . .	49
1.3	L'ordre est respecté. . . . .	59
1.4	L'ordre n'est pas respecté. . . . .	59
1.5	Cas d'objets ponctuels : la contrainte d'ordre n'est plus valable. . . . .	60
1.6	Cas d'objets non ponctuels : la validité de la contrainte d'ordre est assurée. . . . .	60
1.7	Cas d'objets non ponctuels : la contrainte d'ordre n'est pas applicable. . . . .	61
2.1	Géométrie du stéréoscope linéaire. . . . .	81
2.2	Modèle du sténopé. . . . .	82
2.3	Primitives de type point de contour en vision linéaire. . . . .	83
2.4	Extraction des points de contour. . . . .	85
2.5	Matrice de mise en correspondance. . . . .	88
2.6	Mise en défaut de la contrainte de signe du gradient. . . . .	90

---

2.7	Fonction de mesure de compatibilité pour différentes valeur de $\alpha$ . . . . .	95
2.8	Réseau de neurones entièrement connecté. . . . .	97
2.9	Réseau de neurones de Hopfield avec sa matrice des poids de connexions et son vecteur d'entrées externes. . . . .	98
2.10	Réseau de neurones de Hopfield représentant tous les couples possibles. . .	100
2.11	Réseau de neurones de Hopfield ne représentant que les couples candidats à l'appariement. . . . .	101
2.12	Fonction de transfert pour différentes valeurs de $\lambda$ . . . . .	103
2.13	Stéréoscope installé sur le véhicule. . . . .	108
2.14	Séquence stéréoscopique "Piéton". . . . .	109
2.15	Description de la scène "Piéton". . . . .	110
2.16	Image de reconstruction de la séquence "Piéton". . . . .	111
2.17	Séquence stéréoscopique "Approche d'un véhicule". . . . .	113
2.18	Description de la scène "Approche d'un véhicule". Les flèches indiquent les déplacements relatifs des véhicules par rapport au véhicule prototype. . . .	114
2.19	Image de reconstruction de la séquence "Approche d'un véhicule". . . . .	115
3.1	Exemple de matrice de mise en correspondance d'un couple d'images-ligne gauche et droite avec $N_G = 15$ et $N_D = 17$ . Seuls les couples qui vérifient les contraintes locales sont représentés. . . . .	123
3.2	Chromosome binaire. . . . .	123
3.3	Schéma de l'algorithme génétique. . . . .	126
3.4	Séquence "Piéton" : Image de reconstruction obtenue par l'AGB. . . . .	131
3.5	Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par l'AGB. . . . .	131
3.6	Séquence "Piéton" : Image de reconstruction obtenue par l'approche neuronale. . . . .	132
3.7	Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par l'approche neuronale. . . . .	132

---

3.8	Exemple de matrice de mise en correspondance d'un couple d'images-ligne gauche et droite avec $N_G = 15$ et $N_D = 17$ . Seuls les couples qui respectent les contraintes locales sont représentés. . . . .	134
3.9	Chromosome entier. . . . .	136
3.10	Exemple de matrice de mise en correspondance d'un couple d'images-ligne gauche et droite avec $N_G = 6$ et $N_D = 5$ . Seuls les couples vérifiant les contraintes locales sont représentés. . . . .	137
3.11	Exemple d'un chromosome généré par le codage entier : $i \in T_{max} = \{1, 2, 3, 4, 5, 6\}$ et $C_i \in \{0\} \cup T_{min}$ avec $T_{min} = \{1, 2, 3, 4, 5\}$ . . . . .	137
3.12	Chromosome entier et le chromosome binaire correspondant. . . . .	138
3.13	Exemple de matrice de mise en correspondance. . . . .	139
3.14	Exemple de chromosome entier. . . . .	139
3.15	Forme matricielle du chromosome entier de la figure 3.14. . . . .	140
3.16	Evolution de la fonction d'évaluation du meilleur chromosome par l'AGB, avec $Taille_{pop} = 100$ et $N_{gen} = 300$ pour le couple d'images-ligne de rang 135 de la séquence "Piéton". . . . .	142
3.17	Evolution de la fonction d'évaluation du meilleur chromosome par l'AGH, avec $Taille_{pop} = 100$ et $N_{gen} = 300$ pour le couple d'images-ligne de rang 135 de la séquence "Piéton". . . . .	142
3.18	Evolution de la fonction d'évaluation du meilleur chromosome par l'AGB, avec $Taille_{pop} = 200$ et $N_{gen} = 400$ pour le couple d'images-ligne de rang 135 de la séquence "Piéton". . . . .	143
3.19	Séquence "Piéton" : Image de reconstruction obtenue par l'AGE. . . . .	147
3.20	Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par l'AGE. . . . .	147
4.1	Schéma de mise en correspondance au premier niveau de recherche. . . . .	153
4.2	Couples candidats à l'appariement au premier niveau de recherche. . . . .	153
4.3	Schémas de mise en correspondance au deuxième niveau de recherche. . . . .	154

---

4.4	Arbre de recherche de la stratégie hiérarchique. Les couples soulignés représentent les couples de points de contour appariés. . . . .	163
4.5	Séquence "Piéton" : Image de reconstruction obtenue par application hiérarchique de la méthode neuronale. . . . .	164
4.6	Séquence "Piéton" : Image de reconstruction obtenue par application de la méthode neuronale. . . . .	165
4.7	Séquence "Piéton" : Image de reconstruction obtenue par application hiérarchique de l'algorithme génétique entier. . . . .	165
4.8	Séquence "Piéton" : Image de reconstruction obtenue par application de l'algorithme génétique entier. . . . .	165
4.9	Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par application hiérarchique de la méthode neuronale. . . . .	167
4.10	Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par application de la méthode neuronale. . . . .	167
4.11	Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par application hiérarchique de l'algorithme génétique entier. . . . .	168
4.12	Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par application de l'algorithme génétique entier. . . . .	168
4.13	Schéma de mise en correspondance. . . . .	169
4.14	Détermination des couples électeurs d'un couple candidat. . . . .	170
4.15	Séquence "Piéton" : Image de reconstruction obtenue par la méthode de vote. . . . .	173
4.16	Séquence "Piéton" : Image de reconstruction obtenue par la méthode neuronale. . . . .	173
4.17	Séquence "Piéton" : Image de reconstruction obtenue par l'algorithme génétique entier. . . . .	174
4.18	Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par la méthode de vote. . . . .	174
4.19	Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par la méthode neuronale. . . . .	174

---

4.20 Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par l'algorithme génétique entier. . . . .	175
--	-----

---

# Introduction

## Historique

Je commence cette introduction par donner un bref historique de ma carrière universitaire depuis la thèse de doctorat.

### *Thèse de doctorat*

Après un DEA effectué à l'Institut de Recherche en Informatique de Toulouse de l'Université Paul Sabatier, équipe "Traitement et Compréhension d'Images", dirigée par le professeur S. Castan, j'ai débuté ma carrière de chercheur en 1993 par la préparation d'une thèse de doctorat au Centre d'Automatique de Lille de l'Université des Sciences et Technologies de Lille (USTL), équipe "Image et Décision", dirigée par le professeur J.-G. Postaire. Le travail qui m'a été confié consistait à poursuivre les travaux réalisés au sein de cette équipe qui participait au programme PROMETHEUS (Programme for a European Traffic with Highest Efficiency and Unprecedented Safety) pour assurer la détection d'obstacles à l'avant d'un véhicule routier par stéréovision linéaire. J'ai alors poursuivi ces travaux en développant une nouvelle approche pour la stéréovision, basée sur l'utilisation des réseaux de neurones de Hopfield. Parallèlement à cette activité de recherche, j'ai tenu à participer, en tant qu'enseignant vacataire, aux activités pédagogiques de plusieurs établissements universitaires de la Région Nord-Pas-de-Calais.

### *Séjour aux Etats-Unis*

A l'issue de ma thèse, soutenue en février 1997, je me suis lancé dans la recherche d'un post-doc aux Etats-Unis pour enrichir mon expérience et découvrir les conditions de travail outre-atlantique. J'ai alors rejoint le GRASP Lab. (General Robotics and Active Sensory Perception Laboratory) de l'Université de Pennsylvanie pour travailler au sein de l'équipe du professeur R. Bajcsy dans le cadre d'un projet qui concernait la télé-présence. Le travail qui m'a été confié portait sur la mise en oeuvre d'un système stéréoscopique trinoculaire sur une architecture parallèle.

### *ATER à l'USTL*

A l'issue d'un séjour de trois mois aux Etats-Unis et après une courte période de

---

réflexion sur les perspectives à donner à mes activités professionnelles, je suis retourné en France en juillet 1998 pour creuser le chemin d'une carrière d'enseignant-chercheur dans le système universitaire Français. Grâce au soutien de l'équipe "Scènes Dynamiques" du Laboratoire I3D (Image, Interaction et Ingénierie de la Décision), j'ai obtenu en octobre 1998 un emploi d'Attaché Temporaire d'Enseignement et de Recherche (ATER) à temps complet à l'UFR d'IEEA. L'équipe "Scènes Dynamiques", dirigée par J.-G. Postaire, n'était autre que l'équipe "Image et Décision" qui a changé de nom quand le Centre d'Automatique de Lille est devenu en 1998 le "Laboratoire I3D". Cette opportunité a été très avantageuse pour moi car elle m'a permis de continuer mes travaux de recherche et de les développer au sein de la même équipe. Elle m'a conduit à proposer un sujet de DEA à H. Issa que j'ai co-encadré avec J.-G. Postaire. Le bilan de cette première expérience en tant qu'ATER était très positif tant au niveau de la recherche qu'au niveau des enseignements que j'ai pu réaliser. J'ai pu obtenir également en avril 1999 ma qualification aux fonctions de Maître de Conférences dans les sections 27 (Informatique) et 61 (Automatique, Génie Informatique et Traitement du Signal). Au cours de la campagne de recrutement de 1999, il n'y avait qu'un seul poste de Maître de Conférences section 61 au sein de l'UFR d'IEEA de l'USTL. Mais on était quatre candidats locaux plus, bien entendu, les candidats externes. Malgré le soutien de l'équipe "Scènes Dynamiques", la commission de spécialistes a décidé de privilégier les candidatures externes à l'USTL. A l'issue de cette campagne, je n'ai pas pu obtenir un poste de Maître de Conférences, mais les résultats de classement ont été plutôt encourageants : 2<sup>ème</sup> à l'IUT du Havre sur un poste 61<sup>ème</sup> section, 4<sup>ème</sup> à l'IUT A de Lyon sur un poste 27<sup>ème</sup> section et 5<sup>ème</sup> à l'INSA de Rennes sur un poste 61<sup>me</sup> section. Ces résultats m'ont poussé à persévérer et à me lancer dans la campagne de recrutement d'ATER. Cherchant à explorer d'autres opportunités, j'ai élargi le champ de mes candidatures à d'autres universités telles que l'Université de Nice et l'Université de Bretagne Occidentale (UBO) qui m'ont répondu favorablement. L'Université de Nice n'offrait qu'un poste à mi-temps. J'ai choisi donc de partir à Brest pour occuper un poste d'ATER à temps complet au Département Informatique de l'UBO. Parallèlement à ce choix, j'ai tenu à garder un lien fort avec l'équipe "Scènes Dynamiques" au sein de laquelle j'ai appris beaucoup sur la recherche et le métier d'enseignant-chercheur. C'est ainsi

---

que je m'étais engagé à poursuivre le co-encadrement de H. Issa dans le cadre de sa thèse de doctorat.

### ***ATER à l'UBO***

Dès mon arrivée en octobre 1999 à l'UBO, j'ai proposé à l'équipe "Informatique", dirigée par le professeur L. Marcé, ma participation au développement d'un projet sur l'architecture logicielle pour la robotique mobile et téléopérée en mettant à disposition mes compétences en perception de l'environnement. La décision de mettre fin à ce projet et la courte période pour évoluer dans le domaine de l'ingénierie logicielle des systèmes qui dominait les activités de recherche de cette équipe ne m'ont pas permis de m'y intégrer. Parallèlement à mes activités d'enseignement à l'UBO, j'ai continué à co-encadrer les travaux de H. Issa qui préparait sa thèse au sein de l'équipe "Scènes Dynamiques". J'ai passé ainsi sept mois entre Brest, essentiellement pour l'enseignement, et Lille, pour la vie familiale et la recherche. Cette situation a, sans aucun doute, joué en ma défaveur pendant la campagne de recrutement des Maîtres de Conférences de 2000 qui n'a pas été concluante : 2<sup>ème</sup> à l'ENIB de Bourges sur un poste 61<sup>me</sup> section, 3<sup>ème</sup> à l'Université de La Rochelle sur un poste 61<sup>ème</sup> section.

### ***Chercheur contractuel à l'USTL***

A l'issue de mon poste d'ATER à Brest, J.-G. Postaire m'a proposé de retourner au sein de l'équipe "Scènes Dynamiques" en tant que chercheur contractuel dans le cadre du programme STATUE (Système Télématique destiné à l'Accroissement de la sûreté des Transports guidés Urbains et à l'Emergence de nouvelles aides à l'exploitation) pour travailler sur la perception de l'environnement à l'avant des systèmes de transports guidés automatisés et par la même occasion poursuivre mon co-encadrement de la thèse de H. Issa. J'ai accepté cette offre sans aucune hésitation pour continuer à évoluer en suivant le chemin que je m'étais tracé tout au long de ces années au sein de l'équipe de J.-G. Postaire. Cet enchaînement logique dans mes activités de recherche au sein de l'équipe "Scènes Dynamiques" m'a permis de présenter une candidature très soutenue au cours de la campagne de recrutement des Maîtres de Conférences de 2001. La participation à

cette campagne a été couronnée d'un grand succès en décrochant deux fois la position décisive, à l'IUT de Bourges sur un poste 61<sup>ème</sup> section et à l'Université de Technologie de Belfort-Montbéliard sur un poste 27<sup>ème</sup> section (aussi 2<sup>ème</sup> à l'Université de Clermont Ferrand sur un poste 61<sup>ème</sup> section et 5<sup>ème</sup> à l'Université de Tours sur un poste 27<sup>ème</sup>).

### *Maître de Conférences à l'UTBM*

Suite aux discussions avec les collègues des équipes accueillantes au sein de ces deux établissements, j'ai choisi l'Université de Technologie de Belfort-Montbéliard (UTBM) en me rattachant au Département Génie Informatique (GI) et au Laboratoire Systèmes et Transports (SeT), équipe "Informatique", dirigée par le professeur A. Koukam. Cette décision était parfaitement cohérente avec les objectifs que je m'étais tracés tout au long de mon parcours. En effet, sur le plan pédagogique, mon projet d'insertion au sein du Département GI consistait à réorganiser et réorienter les enseignements d'une filière, qui traitait du multimédia, vers les domaines de la vision artificielle et de la réalité virtuelle. Sur le plan de la recherche, mon projet d'insertion concernait le développement d'une nouvelle thématique au laboratoire autour du concept du véhicule intelligent et de son intégration dans la ville du futur, tout en créant une nouvelle synergie entre le domaine des systèmes multi-agents, un des domaines de l'équipe "Informatique", et le domaine de l'analyse d'images. Dans ce cadre, et dès mon arrivée au sein de cette équipe, j'ai proposé un sujet de thèse à M. Hariti que j'ai co-encadré avec A. Koukam. Aujourd'hui, après une durée de quatre ans, le bilan est positif. En effet, la nouvelle filière est mise en place avec succès, la thèse de M. Hariti est soutenue en décembre 2004 et le projet de recherche "Véhicule Intelligent pour la Ville du Futur" est en plein développement autour d'un véhicule expérimental, acheté grâce aux soutiens financiers de nombreux organismes, en particulier la Région Franche Comté. Dans le cadre de ce projet, l'équipe "Informatique" accueillera, dès octobre 2005, P. Kmiotek pour préparer une thèse en co-tutelle avec l'Université des Sciences et Technologies de Cracovie (Pologne). Cette thèse a reçu le soutien financier de l'ambassade de France en Pologne.

Cette réorientation géographique de ma carrière ne m'a pas empêché de continuer à collaborer avec l'équipe "Scène Dynamiques" qui s'intitule actuellement "Vision et Image"

---

suite à la création du Laboratoire d'Automatique, Génie Informatique et Signal (LAGIS) en 2004. Cette continuation a permis de mener à terme la thèse de H. Issa, soutenue en février 2004. Actuellement, je participe au co-encadrement de la thèse de T. Yahiaoui, débutée en septembre 2003, avec F. Cabestaing, Maître de Conférences HDR au sein de l'équipe "Vision et Image", et L. Khoudour, Ingénieur de Recherche au Laboratoire d'Electronique, Ondes et Signaux pour les Transports (LEOST) de l'Institut National de Recherche sur les Transports et leur sécurité (INRETS) de Lille. Cette thèse, qui s'inscrit dans le cadre d'un contrat CIFRE avec la Régie Autonome des Transports de Paris (RATP), concerne l'étude de faisabilité d'un système stéréoscopique pour le comptage de personnes entrant et sortant d'un bus.

### *Passer mon HDR*

Passer une Habilitation à Diriger des Recherches (HDR) est, certes, une étape dans une carrière d'enseignant-chercheur. Plus qu'un objectif, cette étape constitue un engagement. Tout d'abord vis-à-vis du métier d'enseignant-chercheur lui-même, puis vis-à-vis de l'équipe au sein de laquelle l'enseignant-chercheur développe sa démarche de recherche. Il ne faut en aucun cas que cet engagement devienne une ambition personnelle car, en réalité, une HDR constitue un objectif d'équipe dans lequel chacun doit trouver son intérêt. La démarche de recherche du candidat doit incarner tous les ingrédients nécessaires pour concrétiser cet intérêt. Au delà de la démarche scientifique du candidat, ces ingrédients doivent composer un équilibre entre une démarche humaine, un partage de responsabilités et un esprit de concurrence saine. Ceci donne une réponse à la question "Pourquoi passe-t-on une HDR ?". Vient ensuite la question "Quand passe-t-on une HDR". A mon sens, une partie de la réponse à cette question se trouve dans celle que j'ai donnée à la première. La notion de temps ne doit pas agacer le candidat. En effet, quand on est conscient de l'engagement à passer une HDR au sein d'une équipe, cette HDR doit être vue comme un résultat émergeant de l'équipe. Dans mon cas, j'ai eu la chance de conduire l'émergence de mon HDR au sein de deux équipes qui ont deux cultures scientifiques différentes, ce qui m'a permis de mieux me situer par rapport à deux démarches de recherche différentes dans la forme, mais identiques sur le fond.

---

Je finis ce bref historique par évoquer les raisons du choix de l'USTL pour soutenir mon HDR. La première est que l'UTBM n'est pas habilitée à délivrer le diplôme de l'HDR. La deuxième raison est que l'USTL, et en particulier l'équipe "Vision et Image" du Laboratoire LAGIS, a occupé et occupe encore une place prépondérante dans ma carrière de chercheur.

## Contexte et contributions

Depuis quelques années, la vision artificielle a fait des progrès considérables. Le plus remarquable d'entre eux est la possibilité de percevoir la profondeur des objets d'une scène à partir d'images vidéo. Une des approches les plus utilisées dans ce domaine est la vision stéréoscopique. Le système stéréoscopique le plus courant et le plus simple n'utilise que deux images prises sous des angles de vue différents. Il s'agit de la stéréovision binoculaire. Comme chez l'homme, la perception de la profondeur par stéréovision binoculaire est principalement assurée par l'exploitation du décalage existant entre les deux images. Ce décalage, appelé disparité stéréoscopique, ne peut être évalué qu'en établissant des correspondances entre les deux images.

### *Mes premiers travaux de recherche au Centre d'Automatique de Lille*

Mes activités de recherche s'inscrivent principalement dans la perception de l'environnement par analyse de séquences d'images stéréoscopiques, avec comme champ applicatif l'amélioration de la sécurité dans les systèmes de transports terrestres. Mes premiers travaux ont commencé au Centre d'Automatique de Lille de l'USTL, au sein de l'équipe "Image et Décision", qui participait au programme européen PROMETHEUS pour assurer la détection d'obstacles à l'avant d'un véhicule routier par stéréovision linéaire. L'objectif était de poursuivre les travaux de cette équipe, en particulier ceux de J.-L. Bruyelle et J.-C. Burie qui ont effectué leurs thèses dans le cadre de ce programme. Dans sa thèse, J.-L. Bruyelle a présenté la conception ainsi que la réalisation mécanique et électronique d'un stéréoscope à base de caméras linéaires. Par rapport aux caméras matricielles, l'utilisation des caméras linéaires permet une analyse plus rapide des couples

---

d'images stéréoscopiques. Suite à ce travail, J.-C. Burie a présenté dans sa thèse des algorithmes de détection d'obstacles basés sur l'analyse des couples d'images issues du stéréoscope linéaire. Il a proposé une technique d'extraction de points de contour ainsi qu'une méthode séquentielle pour leur mise en correspondance. Quand je suis arrivé au sein de l'équipe "Image et Décision", J.-L. Bruyelle et J.-C. Burie étaient encore là et j'ai pu travailler avec eu durant la dernière année du programme PROMETHEUS. La méthode de mise en correspondance développée par J.-C. Burie ayant montré ses limites, j'ai proposé une nouvelle approche basée sur l'utilisation des réseaux de neurones de Hopfield.

Cette approche consiste à exprimer le problème de mise en correspondance sous la forme d'un problème d'optimisation où une fonction objective doit être minimisée par un réseau de neurones de Hopfield. Ce réseau est entièrement connecté et chaque neurone y représente une possibilité de mise en correspondance entre un point de contour de l'image gauche et un point de contour de l'image droite. Représentant les contraintes du problème, la fonction objective est construite de telle sorte que son minimum corresponde aux bons appariements.

Appliquée aux images enregistrées dans le cadre du programme PROMETHEUS, cette approche s'est révélée très prometteuse et a été exploitée au sein de l'équipe qui participait à de nombreux projets liés à la sécurité dans les transports terrestres. En effet, elle a été utilisée dans le cadre du projet "Carrefour Intelligent" où il s'agissait de localiser et suivre des véhicules en utilisant des contours actifs extraits de séquences d'images stéréoscopiques. Ce travail a été réalisé en collaboration avec C. Vieren, Maître de Conférences à l'USTL et membre de l'équipe "Image et Décision". Dans le cadre du même projet, l'approche neuronale a été adaptée et appliquée à la mise en correspondance temporelle afin de suivre les piétons dans un environnement urbain par analyse de séquences d'images monoculaires. Ce travail a été réalisé en collaboration avec P. Vannoorenberghe (actuellement Maître de Conférences à l'Université de Rouen) dans le cadre de sa thèse effectuée au sein de l'équipe "Image et Décision".

*Poursuite des travaux au laboratoire I3D*

---

Les résultats issus de ma thèse ont été également le point de départ d'autres travaux de recherche au sein de l'équipe "Scènes Dynamiques". Ces travaux auxquels j'ai participé ont donné lieu à deux thèses soutenues au sein de l'équipe. La première thèse, réalisée par D. Dooze, a été effectuée dans le cadre du programme STATUE du PREDIT dont l'un des objectifs était d'assurer la détection d'obstacles à l'avant des systèmes de transports guidés et automatisés. Cette thèse a porté sur le développement d'une nouvelle stratégie de détection d'obstacles en utilisant la stéréovision matricielle avec une option linéaire. Dans le cadre de ce programme, je me suis occupé, plus particulièrement, d'adapter l'approche neuronale au problème de la mise en correspondance des images issues d'un nouveau stéréoscope, baptisé stéréoscope bimodal à portée variable. Ce travail a été réalisé en collaboration avec L. Khoudour et S. Ambellouis du Laboratoire LEOST de l'INRETS de Lille. Dans la deuxième thèse, que j'ai co-encadrée avec J.-G. Postaire, H. Issa a développé, en se basant sur les résultats issus de la méthode neuronale, une approche de mise en correspondance basée sur l'utilisation des algorithmes génétiques qui, contrairement aux réseaux de neurones de Hopfield, sont des méthodes de recherche globale. Nous avons tout d'abord utilisé un algorithme génétique basé sur un codage binaire classique. Nous avons rapidement constaté son inefficacité tant au niveau de la qualité des appariements qu'au niveau du temps de convergence des algorithmes. Nous avons alors proposé un nouveau codage permettant à la fois une meilleure exploration de l'espace des solutions et une accélération des algorithmes.

Dans sa thèse, réalisée au sein de l'équipe "Image et Décision" dans le cadre du programme PROMETHEUS, J.-C. Burie (actuellement Maître de Conférences à l'Université de La Rochelle) a développé une stratégie de recherche hiérarchique pour la mise en correspondance des images stéréoscopiques. Nous avons eu (J.-C. Burie et moi-même) déjà l'occasion de travailler ensemble durant la dernière année de sa thèse pour poursuivre les travaux de recherche qui ont été réalisés dans le cadre de ce programme. Récemment, dans le cadre d'un travail commun, nous avons constaté que cette stratégie de recherche permet de décomposer le problème de mise en correspondance en plusieurs sous-problèmes de combinatoires réduites. Il était donc intéressant d'adapter et d'appliquer cette straté-

---

gie pour diminuer le temps des traitements de nos procédures de mise en correspondance, basées sur les réseaux de neurones de Hopfield et les algorithmes génétiques. En effet, le temps de calcul de ces deux techniques dépend fortement de la combinatoire du problème de mise en correspondance. Le principe de cette stratégie de recherche est de mettre en correspondance les points de contour à plusieurs niveaux de recherches, en commençant par les plus significatifs, c'est à dire ceux ayant les plus fortes amplitudes, pour terminer avec les moins significatifs, c'est à dire ceux ayant de faibles amplitudes. A chaque niveau de recherche, les points de contour les plus significatifs sont sélectionnés puis appariés. Les couples de points de contour obtenus sont ensuite utilisés comme des couples de référence pour mettre en correspondance les points de contour les plus significatifs au niveau suivant.

### *Mes travaux de recherche au laboratoire SeT*

Dès ma nomination à l'UTBM en tant que Maître de Conférences en septembre 2001, j'ai proposé à l'équipe "Informatique" du Laboratoire SeT de mener une nouvelle activité de recherche au sein du laboratoire autour du projet "Véhicule Intelligent pour la Ville du Futur". Ce projet, qui a été approuvé par l'équipe directrice du Laboratoire SeT, est soutenu par plusieurs organismes, en particulier la Région Franche Comté. Sur le plan scientifique, ce projet met en évidence deux grandes thématiques : la perception de l'environnement et la navigation autonome. Ces thématiques sont traitées en collaboration avec le Laboratoire LORIA-INRIA de Nancy, équipe "Machine Autonome et Intelligente", dirigée par F. Charpillet.

Dans le cadre de ce projet, j'ai proposé un sujet de thèse en lien direct avec mes travaux de recherche précédents. Un des objectifs de cette thèse était de développer une procédure de mise en correspondance plus rapide que les deux premières (basées sur les réseaux de neurones de Hopfield et les algorithmes génétiques). La stratégie de recherche hiérarchique multi-niveaux que nous avons développée avec J.-C. Burie permet un gain de temps très substantiel au niveau de l'analyse des couples d'images stéréoscopiques. Cependant, elle est confrontée à deux limitations. La première est liée à la décomposition du problème qui conduit à une perte de l'aspect global de la recherche des appariements.

---

La deuxième limitation résulte de la phase de sélection des points de contour à apparier à chaque niveau de recherche. En effet, à un niveau de recherche donné, un point de contour peut être sélectionné dans une image alors que son correspondant dans l'autre image ne l'est pas. Face à ces difficultés, nous avons proposé dans le cadre de la thèse de M. Hariti, que j'ai co-encadrée avec A. Koukam, une nouvelle approche de mise en correspondance stéréoscopique. Elle consiste à affecter un score à chaque couple de primitives suivant une stratégie de vote, régie par les contraintes du problème. Grâce à sa simplicité de mise en oeuvre, cette méthode permet une mise en correspondance rapide tout en fournissant des résultats très comparables à ceux obtenus par les approches neuronale génétique.

Le projet "Véhicule Intelligent pour la Ville du Futur" m'a conduit à travailler sur la navigation autonome d'un véhicule à partir de la perception de son environnement. Cette nouvelle thématique porte sur la localisation de véhicules autonomes et la planification de leurs mouvements. En plus de la nouvelle orientation qu'elle donne à mes activités de recherche, cette thématique me permet d'exploiter mes travaux sur la perception de l'environnement par stéréovision. Les premiers résultats ont été présentés à Nancy, les 17 et 18 juin, sous la forme d'un démonstrateur à l'occasion de l'année des lumières dans le cadre du programme MobiVIP (Véhicules Individuels Publics pour la Mobilité en centre ville) soutenu par le PREDIT. Le Laboratoire SeT a participé à cette manifestation en tant que partenaire de l'équipe MAIA qui fait partie du consortium du programme MobiVIP. Mes travaux sur la navigation autonome en sont à leur début et seront décrits en termes de perspectives. Dès octobre 2005, l'équipe "Informatique" accueillera P. Kmiotek dans le cadre d'une thèse en co-tutelle avec l'Université des Sciences et Technologies de Cracovie (Pologne). Cette thèse, qui a bénéficié du soutien financier de l'ambassade de France en Pologne, portera sur la perception par fusion multi-capteurs et sur la navigation autonome.

Par ailleurs, depuis mon arrivée au sein de l'équipe "Informatique" du laboratoire SeT, j'ai créé une nouvelle synergie en associant le domaine des systèmes multi-agents, un des domaines de compétence de cette équipe, et le domaine de l'analyse d'images. Il s'agit de développer des modèles multi-agents pour résoudre des problèmes en analyse

d'images tels que la détection de contours, la segmentation, la modélisation d'objets par les modèles de contours actifs et la mise en correspondance. Ces travaux en sont à leur début et seront décrits en termes de perspectives.

### *Mes collaborations avec le laboratoire LAGIS*

Malgré la situation géographique, je continue toujours à collaborer avec l'équipe "Scène Dynamiques", connue actuellement sous le nom "Vision et Image" au sein du Laboratoire LAGIS de l'USTL, et le Laboratoire LEOST de l'INRETS de Lille. Nous travaillons à présent sur l'étude de faisabilité de la stéréovision pour le comptage de passagers entrant et sortant d'un bus. Ce travail, mené dans le cadre d'un contrat CIFRE avec la RATP, fait l'objet de la thèse de T. Yahiaoui, débutée en septembre 2003 et co-encadrée par F. Cabestaing, Maître de Conférences HDR à l'USTL et membre de l'équipe "Vision et Image", L. Khoudour, Ingénieur de Recherche au Laboratoire LEOST de l'INRETS, et moi-même.

## Organisation du mémoire

Bien qu'il ne porte qu'une seule signature, ce mémoire présente les travaux auxquels j'ai eu le plaisir de participer en collaboration avec plusieurs collègues, doctorants et enseignants-chercheurs. De manière traditionnelle, je commence par présenter un Curriculum Vitae qui résume mes activités personnelles, en termes de participation aux activités d'enseignement, d'encadrement de chercheurs, de publications et de projets. Les principaux résultats de mes activités de recherche sont ensuite décrits en quatre chapitres.

Après un état de l'art sur la stéréovision, le deuxième et le troisième chapitres présentent respectivement les approches de mise en correspondance par réseaux de neurones de Hopfield et par algorithmes génétiques. Le quatrième chapitre présente la stratégie de recherche hiérarchique et la technique de vote permettant de réaliser une mise en correspondance stéréoscopique rapide.

Une conclusion clôture ce mémoire en soulignant les perspectives selon lesquelles je

compte poursuivre mes travaux de recherche.

---

# Notice individuelle

## Yassine Ruichek

36 ans, né le 10 septembre 1969 à Marrakech (Maroc)

Marié, 2 enfants (nés en janvier 2001 et en janvier 2004)

Maître de Conférences à l'Université de Technologie de Belfort-Montbéliard

Enseignement : Département Génie Informatique

Recherche : Laboratoire Systèmes et Transports

### Adresse professionnelle

Laboratoire Systèmes et Transports

Université de Technologie de Belfort-Montbéliard

90010 Belfort Cedex

Tél : 03 84 58 34 41 (direct) – 03 84 58 33 19 (secrétariat)

Fax : 03 84 58 33 42

Courriel : yassine.ruichek@utbm.fr

### Titres universitaires

**1989** DEUG mathématiques physiques (Passable)

Université Cady Ayad, Marrakech (Maroc)

**1991** Licence Informatique (Passable)

Université Paul Sabatier, Toulouse

**1992** Maîtrise Ingénierie Mathématique (A. Bien)

Université Paul Sabatier, Toulouse

**1993** DEA Interaction Homme Système Multimodale (A. Bien)

Université Paul Sabatier, Toulouse

**1997** Doctorat Automatique et Informatique Industrielle (Très Honorable)

Université des Sciences et Technologies de Lille

---

## Doctorat

Stéréovision linéaire par réseau de neurones de Hopfield – Application à la détection d'obstacles à l'avant des véhicules routiers.

Soutenue : le 04 février 1997 à l'Université des Sciences et Technologies de Lille

Spécialité : Productique – Automatique et Informatique Industrielle

Mention : Très Honorable

Membres du jury :

- Pierre Vidal (Président)
- Isabelle Magnin (Rapporteur)
- Alain Faure (Rapporteur)
- Mohamed Benjelloun (Examineur)
- Mohamed Karim (Examineur)
- Christian Louis (Examineur)
- Denis Hamad (Examineur)
- Jack-Gérard Postaire (Directeur de thèse)

## Activités professionnelles

- |                    |   |
|--------------------|---|
| <b>1997</b>        | Chercheur invité (Postdoc)<br>Université de Pennsylvanie (GRASP Lab.), Etats-Unis<br>septembre 1997 – décembre 1997               |
| <b>1998 – 1999</b> | Attaché Temporaire d'Enseignement et de Recherche<br>Université des Sciences et Technologies de Lille<br>octobre 1998 – août 1999 |
| <b>1999 – 2000</b> | Attaché Temporaire d'Enseignement et de Recherche<br>Université de Bretagne Occidentale<br>octobre 1999 – août 2000               |
| <b>2000 – 2001</b> | Chercheur contractuel<br>Université des Sciences et Technologies de Lille (laboratoire I3D)<br>septembre 2000 – août 2001         |

**Depuis 2001** Maître de Conférences

Université de Technologie de Belfort-Montbéliard

---

# Activités d'enseignement

## Parcours

De 1993 à 1997, j'ai été enseignant vacataire dans plusieurs établissements universitaires de la région Nord-Pas-de-Calais. J'ai ensuite enseigné en qualité d'Attaché Temporaire d'Enseignement et de Recherche (ATER) à temps plein à l'Université des Sciences et Technologies de Lille durant la période 1998 – 1999, puis à l'Université de Bretagne Occidentale à Brest durant la période 1999 – 2000. Depuis septembre 2001, j'occupe un poste de Maître de Conférences à l'Université de Technologie de Belfort-Montbéliard.

### En qualité de contractuel : enseignant vacataire et ATER

Durant cette première expérience pédagogique, je suis intervenu dans plusieurs disciplines de l'IEEA (Informatique, Electronique, Electrotechnique et Automatique) en m'adressant à des publics variés (DUT, DEUG, Licence, Maîtrise, DESS, Cycles d'Ingénieurs de l'Institut Agro-Alimentaire de Lille et de l'Ecole Nouvelle d'Ingénieurs en Communication de Lille). Les enseignements que j'ai effectués durant cette période (1993 – 2000) sont résumés sur le tableau récapitulatif ci-dessous. Les volumes horaires sont exprimés en heure TD.

Matière	1 <sup>er</sup> cycle	2 <sup>ème</sup> cycle	3 <sup>ème</sup> cycle
Automatique (TP)		48 heures	
Automatisme (TD et TP)	120 heures	6 heures	
Traitement du signal et de l'image (TD)		12 heures	
Réseaux de neurones (Cours)			10 heures
Mathématiques (TD)	20 heures		
Algorithme et programmation (Cours, TD et TP)	246 heures	32 heures	
Programmation fonctionnelle (TD et TP)	54 heures		
Bases de données (TD et TP)	20 heures		
Bureautique (TP)	12 heures		
<b>TOTAL</b>	<b>472 heures</b>	<b>98 heures</b>	<b>10 heures</b>

---

## En qualité de Maître de Conférences à l'UTBM

### Le cursus universitaire de l'UTBM

L'Université de Technologie de Belfort-Montbéliard (UTBM) est une école d'ingénieurs appartenant au réseau des Universités de Technologie composé de trois établissements : l'UTC (Compiègne), l'UTT (Troyes) et l'UTBM (Belfort-Montbéliard). Le cursus universitaire, structuré en semestres, est de cinq ans : deux ans en tronc commun et trois ans en branche. L'UTBM propose quatre branches : Génie Informatique, Génie des Systèmes de Commande, Génie des Systèmes de Production et Génie Mécanique de Conception. Dans chaque branche, les enseignements sont répartis en deux catégories. La première est constituée d'unités de valeur de base (UVs de connaissances de base) destinées aux élèves-ingénieurs qui sont en première année de branche. A l'issue de cette première expérience en branche, les élèves-ingénieurs partent en stage de technicien pour une période de six mois. De retour de stage, ils attaquent pendant deux semestres la deuxième catégorie d'enseignement constituée d'UVs spécifiques (UVs de compétences spécifiques). Celles-ci sont regroupées en quatre filières : **ILC** (Ingénierie des Logiciels et de la Connaissance, **RT** (Réseaux et Télécommunication), **REM** (Systèmes Temps Réel, Embarqués et Informatique Mobile) et **I2RV** (Image, Interaction et Réalité Virtuelle). Cette dernière filière, mise en place depuis 2003, remplace une ancienne filière qui s'appelait ICI (Image, Communication et Information). Pour compléter leur profil, les élèves-ingénieurs doivent valider des UVs de communication et de culture générale. A l'issue des études de branche, les élèves-ingénieurs repartent en stage de projet de fin d'études pour une période de six mois.

### Bilan général

Depuis ma nomination en septembre 2001 au département Génie Informatique (GI) de l'UTBM, j'interviens à la fois dans des UVs de base et des UVs spécifiques. J'interviens également, depuis 2002, dans une UV de compétence spécifique au département Génie des Systèmes de Commande (GSC) et dans le Master Recherche Mécatronique depuis 2004.

Ma première expérience en enseignement et mes activités de recherche m'ont permis de

mener à bien mes activités pédagogiques au sein de l'UTBM. Dès mon arrivée au département GI, j'ai créé une nouvelle dynamique pour renouveler les orientations pédagogiques d'un certain nombre d'enseignements. C'est ainsi que j'ai mis en place une équipe qui avait comme objectif l'élaboration d'un projet de filière autour de la vision artificielle et de la réalité virtuelle. Ce projet, baptisé I2RV (Image, Interaction et Réalité Virtuelle), a été adopté par le département GI qui m'a désigné comme responsable pédagogique et administratif de cette filière. Cette responsabilité m'a conduit à participer activement à la vie pédagogique au sein du département.

Etant donné que l'UTBM est une école d'ingénieurs, j'ai opté pour une démarche pédagogique orientée recherche et innovation, en particulier pour les enseignements de la filière I2RV étroitement liés à mes activités de recherche. Sachant que les travaux pratiques occupent une partie importante dans la formation d'ingénieur, je favorise toujours la participation des élèves-ingénieurs à des sujets de recherche concrets. Cette démarche permet de rapprocher les élèves-ingénieurs des activités de recherche du laboratoire tout en poursuivant les objectifs pédagogiques des enseignements. En plus de la mise en pratique des acquis fondamentaux, les sujets proposés permettent aux élèves-ingénieurs de se confronter à des problèmes réels d'intégration sur les plates-formes du laboratoire SeT, en particulier les plate-formes "Véhicule Intelligent" et "Réalité Virtuelle" dont les objectifs sont étroitement liés aux enseignements de la filière I2RV. Cette démarche a tendance de temps en temps à sensibiliser les élèves-ingénieurs au domaine de la recherche.

### **Bilan des enseignements**

Les graphiques ci-dessous donnent les répartitions des activités d'enseignement par matière, par niveau, par cours, TD, TP et par type d'enseignement. Un bilan détaillé par matière est présenté dans l'annexe 1.

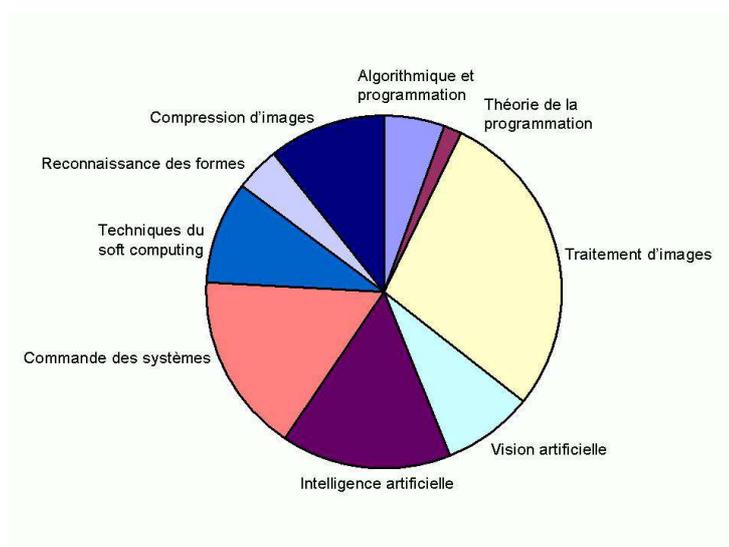


FIG. 1 – Répartition des enseignements par matière.

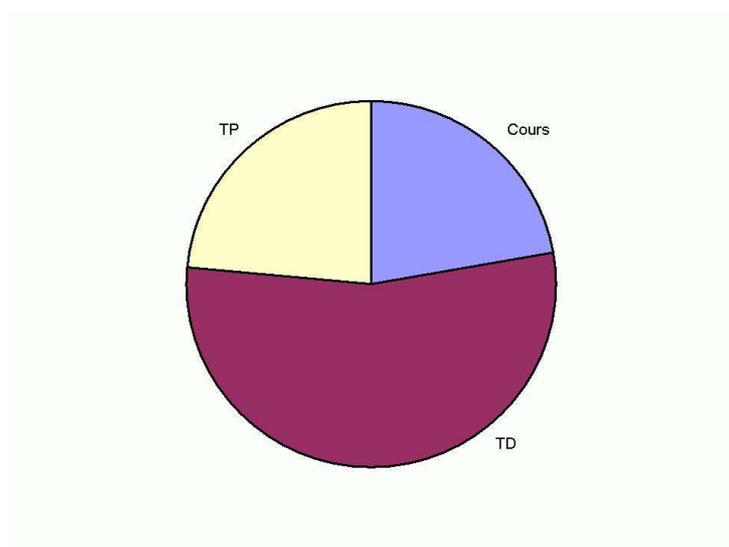


FIG. 2 – Répartition des enseignements par Cours, TD et TP.

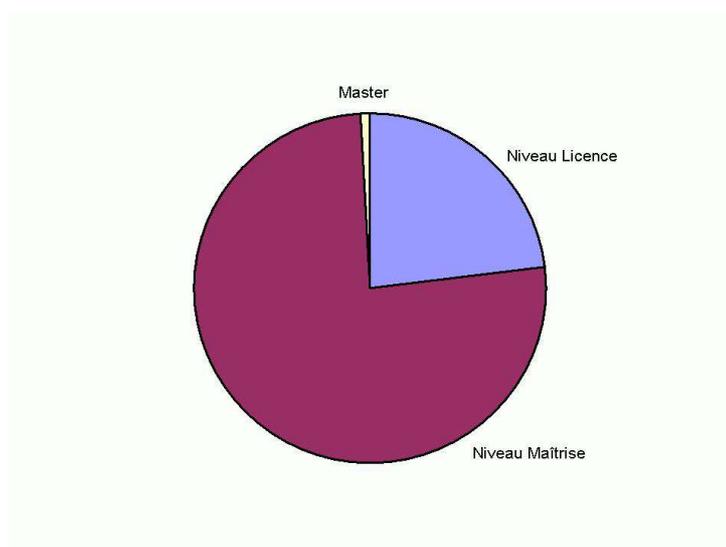


FIG. 3 – Répartition des enseignements par niveau.

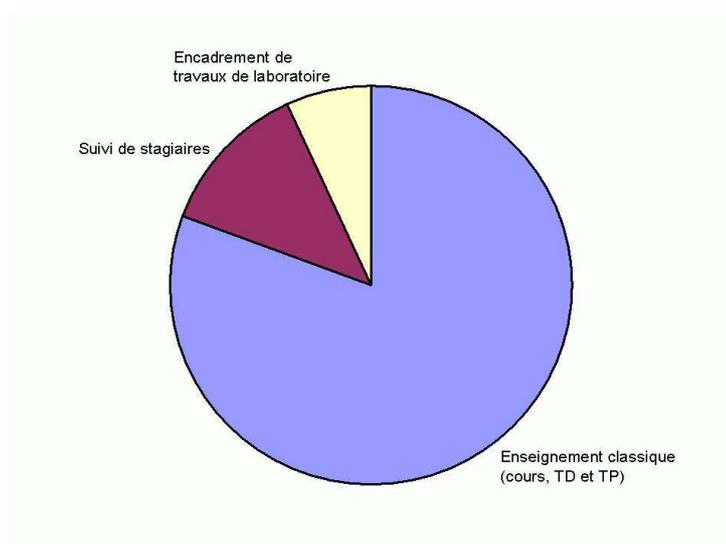


FIG. 4 – Répartition par type d'enseignement.

### Encadrement de stagiaires

En plus de la charge d'enseignement classique (cours, TD et TP), les enseignants-chercheurs de l'UTBM sont amenés à suivre les élèves-ingénieurs qui partent en stage. Ayant à ma charge 6 à 10 stagiaires par an, le suivi consiste à les conseiller en cas de difficulté, les visiter, évaluer leurs rapports puis participer au jury de soutenance. Cette activité de suivi,

qui est d'ordre pédagogique, permet d'être à l'écoute des élèves-ingénieurs pour mieux les accompagner dans leurs premières expériences professionnelles. Elle permet également aux enseignants-chercheurs de prendre connaissance des nouveaux besoins industriels pour mieux orienter les enseignements qu'ils proposent. Par ailleurs, cette activité ouvre les possibilités de partenariat entre l'UTBM et les entreprises pour monter des projets en commun.

### **Encadrement de projets**

Dans le cadre de l'unité de valeur "Travaux de laboratoire", les enseignants-chercheurs de l'UTBM sont invités à proposer aux élèves-ingénieurs des projets d'initiation à la recherche d'une durée de quatre mois. Dans ce cadre, j'ai proposé et encadré plusieurs projets liés à mes activités de recherche sur la perception par vision artificielle. Ces projets sont destinés en particulier aux élèves-ingénieurs de la filière I2RV qui bénéficient d'une formation adaptée pour les développer. Ils traitent des problématiques liées à la thématique "Véhicule Intelligent pour la Ville du Futur" que je mène depuis mon arrivée au laboratoire SeT. La liste suivante donne quelques exemples de projets que j'ai proposés et encadrés :

- Simulation multi-agents d'un véhicule intelligent
- Modélisation d'objets par contours actifs en utilisant une approche multi-agents
- Mise en correspondance d'images stéréoscopiques par un réseau de neurones multicouches
- Détection de contours par une approche multi-agents
- Simulation multi-agents d'un capteur stéréoscopique à deux caméras
- Simulation de la navigation manuelle d'un véhicule sur une plate-forme de réalité virtuelle
- Détection et reconnaissance de balise "p-similaires"

---

## Une expérience pédagogique spécifique : la filière I2RV

Mon projet pédagogique d'insertion en qualité de Maître de Conférences à l'UTBM consistait à prendre en charge la responsabilité d'une filière (existante) au sein du département Génie Informatique, en la réorientant vers les domaines de la vision artificielle et de la réalité virtuelle. Cette filière qui portait le nom ICI (Image, Communication et Information) traitait des sujets liés au développement d'applications multimédias.

A ma nomination en septembre 2001, et en collaboration avec les collègues M. Wack, B. Mignot, O. Lamotte et S. Galland, j'ai commencé par analyser les enseignements de la filière ICI qui étaient répartis sur 5 UVs (nombre total d'UVs autorisé dans une filière). Trois axes ont été mis en évidence. Le premier concernait le traitement et la synthèse d'images. Le deuxième axe regroupait les enseignements en multimédia, tels que la recherche dans les bases de données d'images, les méthodologies de conception et d'analyse de produits multimédias. Le dernier axe portait sur les outils de développement d'applications réparties en termes d'architectures distribuées et d'outils pour la programmation parallèle.

Après deux ans de concertation et en s'inspirant des thématiques liées à mes activités de recherche, j'ai élaboré un projet de filière avec une nouvelle organisation et de nouveaux enseignements. Ce projet de filière, baptisé I2RV (Image, Interaction et Réalité Virtuelle), s'est développé autour du concept de l'image en tant que source et produit d'information pour l'aide à la décision. En tant que source d'information, il s'agit de traiter et analyser l'image pour extraire les informations utiles à son interprétation. En tant que produit d'information, l'objectif est de construire des mondes 2D ou 3D, fixes ou animées, à partir de données simulées.

La nouvelle organisation proposée est structurée autour de trois axes. Le premier concerne la vision artificielle. Il comprend la vision bas niveau, c'est à dire le traitement d'images, et la vision au niveau intermédiaire, en particulier la stéréovision, la reconstruction 3D et la reconnaissance des formes. Le deuxième axe porte sur la réalité virtuelle. Il traite la synthèse d'images, la 3D temps réel et les techniques de création de mondes virtuels

3D en termes de modélisation et de visualisation. Le troisième axe couvre les notions d'interaction et de communication à travers la conception et la création d'applications multimodales en ligne.

Soumis au bureau de département Génie Informatique, puis à la direction de l'UTBM, le projet de la filière I2RV a été adopté en septembre 2003. J'ai été ensuite désigné, par élection au sein du bureau de département Génie Informatique, comme responsable pédagogique et administratif de cette filière à partir de septembre 2003.

Ce projet de filière m'a permis de mener une réflexion à la fois scientifique et pédagogique. Mes activités de recherche ont été une source d'inspiration durant toute la période d'élaboration de ce projet. Cependant, si la filière I2RV a vu le jour avec une réussite et un succès indéniables, c'est, sans aucun doute, grâce aux collègues qui ont fait et font toujours un travail merveilleux.

## Responsabilités pédagogiques

- Responsable pédagogique et administratif de la filière I2RV, département Génie Informatique, depuis septembre 2003
- Membre élu du bureau de département Génie Informatique, depuis septembre 2003
- Membre du jury de suivi des études du département Génie Informatique, depuis septembre 2003
- Membre de la commission de validation des sujets de "travaux de laboratoire", département Génie Informatique, depuis septembre 2003
- Responsable pédagogique de l'UV "Traitement et analyse d'images" (filiales ICI, puis I2RV), département Génie Informatique, depuis septembre 2002
- Responsable pédagogique de l'UV "Reconnaissance des formes" (filiale I2RV), département Génie Informatique, depuis septembre 2004
- Responsable pédagogique de l'UV "Intelligence artificielle en automatique" (filiale Commande des Systèmes), département Génie des Systèmes de Commande, depuis septembre 2002
- Responsable pédagogique de l'UV "Traitement avancé de l'information" (filiale I2RV), département Génie Informatique, en 2003.

---

## Activités de recherche

## Laboratoires de rattachement

Comme je l'ai souligné dans l'historique de ma carrière universitaire au début de l'introduction du manuscrit, j'ai travaillé avec deux équipes ayant deux cultures scientifiques différentes.

La première équipe, au sein de laquelle j'ai commencé ma carrière de chercheur, est composée d'un groupe d'automaticiens de l'USTL, dirigé par J.-G. Postaire. Cette équipe a appartenu successivement au Centre d'Automatique de Lille, au Laboratoire I3D (Interaction, Image et Ingénierie de la Décision) puis au Laboratoire d'Automatique, Génie Informatique et Signal (LAGIS). Officiellement, je ne suis plus membre de l'équipe de J.-G. Postaire depuis ma nomination en tant que Maître de Conférences à l'UTBM en 2001. Mais j'ai continué à travailler sans arrêt avec cette équipe qui m'a toujours soutenu. D'ailleurs, une grande partie des travaux présentés dans ce mémoire a été effectuée au sein de cette équipe.

En 1993, j'ai commencé mes travaux de thèse au sein du Centre d'Automatique de Lille.

### **1993 – 1997 : Centre d'Automatique de Lille (UPRES EA 1022)**

Directeur : P. Vidal

Equipe "Image et Décision" – Responsable : J.-G. Postaire

En 1998, est né le Laboratoire I3D (Interaction, Image et Ingénierie de la Décision) à l'issue de la fusion du Centre d'Automatique de Lille et du Laboratoire de Mesures Automatiques. En janvier 2002, le Laboratoire I3D a été associé au Département STIC du CNRS en tant que FRE 2497.

### **1998 – 2003 : Laboratoire I3D (FRE CNRS 2497)**

Directeur : C. Vasseur

Equipe "Scènes Dynamiques" – Responsable : J.-G. Postaire

En janvier 2004, le Laboratoire I3D et le Laboratoire LAIL (Laboratoire d'Automatique et Informatique Industrielle de Lille) ont fusionné pour donner naissance au Laboratoire LAGIS (Laboratoire d'Automatique, Génie Informatique et Signal). Ce Laboratoire commun à l'Ecole Centrale de Lille et à l'Université des Sciences et Technologies de Lille constitue une Unité Mixte de Recherche du CNRS.

### **2004 Laboratoire LAGIS (UMR CNRS 8146)**

Directeur : P. Vanheeghe

Equipe "Vision et Image" – Responsable : J.-G. Postaire

La deuxième équipe que j'ai rejoint en 2001 dès ma nomination en tant que Maître de Conférences à l'UTBM est composée d'un groupe d'informaticiens, dirigé par A. Koukam au sein du Laboratoire SeT (Systèmes et Transports).

### **Depuis 2001 : Laboratoire SeT**

Directeur : A. El Moudni

Equipe "Informatique" – Responsable : A. Koukam

## Thématiques et projets de recherche

### Thématiques de recherche

Dans la section "Contexte et contribution" de l'introduction de ce mémoire, j'ai exposé un résumé de l'ensemble des travaux auxquels j'ai participé tout au long de ma carrière universitaire. Ces travaux se sont développés et enchaînés autour de :

1. **L'analyse d'images** : Stéréovision dynamique, Détection et suivi d'objets, Optimisation sous contraintes par réseaux de neurones et algorithmes génétiques
2. **La navigation autonome** : Localisation de véhicules autonomes, Planification du mouvement
3. **La modélisation multi-agents en analyse d'images**
4. **Application à la sécurité dans les systèmes de transports terrestres**

### Projets de recherche actuels

#### 1. **Projet "Véhicule Intelligent pour la Ville du Futur"**

Dès ma nomination à l'UTBM, l'équipe directrice du Laboratoire SeT m'a chargé de développer une nouvelle activité de recherche autour du concept du véhicule intelligent et son intégration dans la ville du futur. Il s'agit de développer des systèmes embarqués pour la perception et la navigation autonome d'un véhicule dans un environnement urbain. Une fois le projet a été élaboré, il a fallu solliciter le soutien du monde socio-économique pour acquérir un véhicule expérimental ainsi que les moyens d'instrumentation nécessaires au développement du projet. Actuellement, le Laboratoire SeT dispose d'un véhicule électrique à contrôle automatique de type RobuCab équipé de nombreux capteurs achetés grâce à une subvention de la Région Franche Comté de 45K euros. Le projet a reçu également les soutiens de l'Anvar pour 4,5K euros, de la DRIRE/FEDER pour 22K euros de l'Institut de Transport pour 10K euros et du Département Génie Informatique de l'UTBM pour 8K euros. Le projet implique actuellement trois enseignants chercheurs,

un ingénieur de recherche ainsi que de nombreux étudiants en Master ou projet de fin d'études. Dans le cadre d'une co-tutelle avec l'Université des Sciences et Technologies de Cracovie, le Laboratoire SeT accueillera à partir du mois d'octobre 2005 un doctorant pour travailler sur ce projet. Cette thèse a bénéficié du soutien financier de l'ambassade de France en Pologne. On prévoit également l'accueil d'un deuxième doctorant dans le cadre d'une bourse offerte par la région Franche Comté. La demande de cette bourse a été classée en 2<sup>ème</sup> position par le Conseil Scientifique de l'UTBM. Le projet se développe en partenariat avec le Laboratoire LORIA-INRIA de Nancy, équipe Machine Autonome et Intelligente, qui dispose du même véhicule expérimental que celui du laboratoire SeT. Les deux laboratoires ont présenté à Nancy, les 17 et 18 juin 2005, deux démonstrateurs à l'occasion de l'année des lumières dans le cadre du projet MobiVIP (Véhicules Individuels Publics pour la Mobilité en centre ville) soutenu par le PREDIT. D'autres partenariats sont en cours d'élaboration, notamment avec la société PIMENTIC (Projets Industriels, Mobilité et Nouvelles Technologies de l'Information et de la Communication) basée à Belfort. Avec ce projet, le Laboratoire SeT renforce sa participation au pôle de compétitivité en Alsace – Franche Comté sur le thème "Véhicule du Futur".

## **2. Projet NAPE "Nouvelles Applications de la Perception Embarquée"**

Ce projet, en cours d'élaboration, s'inscrit dans le cadre du réseau RT3 (Réseau Inter-régional de Recherche Technologique pour les Transports Terrestres), Groupe Thématique GT4 (Systèmes Electroniques et Informatiques Embarqués). Il s'agit de développer des systèmes embarqués de perception en intégrant la route comme un système à part entière au service de tous les usagers. Le but de ce projet que je coordonne est d'améliorer la fluidification et la sécurité du trafic de véhicules coopératifs par communication inter-véhicules et/ou véhicule-infrastructure.

## Activités d'encadrement et publications

### Activités d'encadrement

- 3 thèses (2 soutenues et 1 en cours)
- 5 DEA
- 3 stagiaires

Ces activités d'encadrement sont détaillées en annexe 2.

### Publications

- **6 articles dans des revues internationales avec comités de lecture** : Pattern Recognition Letters – Mathematical and Computer Modelling – Real-Time Imaging – Systemics, Cybernetics and Informatics – IEEE Transactions on Intelligent Transportation Systems – Soft Computing
- **3 chapitres de livres** : Computational Imaging and Vision (Springer Ed.) – Advances in Multimedia, Video and Signal Processing Systems (WSEAS Ed.) – Road Vehicle Automation (Wiley Ed.)
- **17 communications dans des congrès internationaux avec actes et comités de lecture**
- **5 communications dans des congrès nationaux avec actes et comités de lecture**
- **2 communications avec actes pour la diffusion de la connaissance**
- **1 article soumis** (International Journal of Image and Vision Computing)

La liste des publications est détaillée en annexe 3.

## Collaborations

### Collaborations institutionnelles

1. Laboratoire LAGIS de l'USTL, équipe "Vision et Image" (J.-G. Postaire et F. Cabestaing)
  - Co-encadrement d'une thèse (H. Issa, soutenue en 2004 à l'USTL) avec J.-G. Postaire
  - Co-encadrement d'une thèse (T. Yahiaoui, débutée en 2003) avec F. Cabestaing et L. Khoudour (Laboratoire LEOST de l'INRETS). Cette thèse s'effectue dans le cadre d'un contrat CIFRE avec la RATP.
2. INRETS, Laboratoire LEOST (L. Khoudour et S. Ambellouis)
  - Programme STATUE du PREDIT (2000 – 2001) : Sous-programme "Détection d'obstacles à l'avant des systèmes de transports guidés automatisés". Cette collaboration s'est effectuée dans le cadre de mon contrat de chercheur contractuel de septembre 2000 à août 2001 au Laboratoire I3D de l'USTL.
3. Laboratoire LORIA-INRIA de Nancy, équipe MAIA (F. Charpillet)
  - Projet commun sur la navigation de véhicules autonomes dans des environnements urbains, depuis 2004.
  - Deux démonstrateurs communs présentés les 17 et 18 juin 2005 à Nancy à l'occasion de l'année des lumières dans le cadre du projet MobiVIP (Véhicules Individuels Publics pour la Mobilité en centre ville) du PREDIT.
4. Institute of Computer science de l'Université des Sciences et Technologies de Cracovie (K. Cetnarowicz)
  - Une thèse en co-tutelle entre l'UTBM et l'Université des Sciences et Technologies de Cracovie (Pologne). P. Kmiotek débutera cette thèse en octobre 2005. Une bourse a été accordée au candidat par l'ambassade de France en Pologne. Le co-encadrement sera assuré par K. Cetnarowicz, A. Koukam et moi-même.

## Collaborations ponctuelles

1. J.-C. Burie (Laboratoire L3I de l'Université de La Rochelle)
  - Travaux communs sur le problème de mise en correspondance d'images stéréoscopiques (Une publication en commun dans la conférence internationale IEEE ICTAI, 2004).
2. C. Vieren (Equipe "Vision et Image", Laboratoire LAGIS de l'USTL)
  - Travaux communs sur la localisation et le suivi de véhicules par mise en correspondance des contours actifs. Ce travail s'inscrivait dans le cadre du projet "Carrefour Intelligent" de l'INRETS (Une publication en commun dans la conférence internationale EANN, 1996).
3. P. Vannoorenberghe (Laboratoire PSI de l'Université de Rouen)
  - Travaux communs sur le suivi de piéton dans un environnement urbain par mise en correspondance temporelle. Ces travaux s'inscrivait dans la cadre du projet "Véhicule Intelligent" de l'INRETS (Une publication en commun dans la conférence internationale EANN, 1996).

## Rayonnement

### Animations scientifiques

- Responsable du projet "Véhicule Intelligent pour la Ville du Futur", soutenu par le Conseil Régional de Franche Comté, depuis 2001.
- Responsable du projet "Système d'Aide à la Conduite par Combinaison de la Perception des Environnements avant et arrière d'un véhicule", soutenu par l'ANVAR (avril 2002 – avril 2003) dans le cadre du contrat No. 02.04.017I JJ d'aide aux jeunes chercheurs pour l'innovation.
- Coordinateur depuis 2004 du projet "Nouvelles Applications de la Perception Embarquée" (NAPE) dans le cadre du réseau RT3, Groupe Thématique GT4 "Systèmes Electroniques et Informatiques Embarqués".
- Membre du comité scientifique (représentant du Laboratoire SeT) du projet européen SURE (A time-oriented model for Sustainable Urban Regeneration), 2002 – 2005.
- Membre du comité technique IEEE Robotic and Automation Society on Intelligent Transportation Systems, depuis 2004.
- Membre de l'editorial board du Journal International Systemics, Cybernetics and Informatics, depuis 2003.
- Membre des comités techniques "Computer vision" et "Neural networks" de l'IASTED, depuis 2003.

### Organisation de conférences et de colloques

- Organisateur et président de la session "Environment Perception for Intelligent Vehicles" dans la conférence internationale IEEE on Networking, Sensing and Control, Taipei, Taiwan, 21–23 mars 2004 (7 communications ont été présentées).
- Membre du comité d'organisation du colloque "Systèmes Embarqués et Transports Terrestres", Université de Technologie de Belfort-Montbéliard, mars 2005.

- Membre du comité d'organisation de la journée technique "Réalité Virtuelle dans l'Industrie et les Transports", Université de Technologie de Belfort-Montbéliard, juin 2004.
- Membre du comité d'organisation de la journée technique "Sécurité dans les Transports Terrestres", Université de Technologie de Belfort-Montbéliard, juin 2003.
- Membre des comités de programmes des conférences IASTED "Neural Networks and Computational Intelligence" (2003 et 2004) et "Computational Intelligence" (2005).
- Membre du comité de programme de la conférence IEEE on Networking, Sensing and Control, 2004.
- Membre du comité scientifique de la conférence ICGST on Graphics, Vision and Image Processing, 2005.

### **Referee d'articles pour des revues internationales**

- Pattern Recognition Letters
- Real-Time Imaging
- Signal Processing
- Soft Computing
- Systemics, Cybernetics and Informatics
- IEEE Trans. on Intelligent Transportation Systems
- Plusieurs conférences internationales

### **Responsabilités collectives**

- Membre élu de la commission de spécialistes de l'UTBM (sections 27 et 61), depuis avril 2003.
- Membre élu du Conseil de Laboratoire SeT, depuis mars 2004.
- Membre de la commission d'étude chargée de la mise en place de la plate-forme de réalité virtuelle immersive, subventionnée par la DRIRE/FEDER, au sein du Laboratoire SeT (2001 – 2003).
- Responsable et animateur des séminaires du Laboratoire SeT (2002 – 2003).

---

---

# CHAPITRE 1

---

Perception par vision stéréoscopique

## 1.1 Introduction

La vision est notre sens le plus puissant. Elle nous fournit toutes sortes d'informations sur les objets qui nous entourent et nous permet de réagir intelligemment et rapidement à notre environnement. Nous avons ainsi la capacité de percevoir les formes, les couleurs et les mouvements des objets. Mais nous sommes aussi capables de structurer notre environnement, d'évaluer les grandeurs et les distances grâce à notre système de vision binoculaire.

La vision est aussi notre sens le plus complexe. Bien que cette faculté semble tellement naturelle qu'elle passe inaperçue, elle est le résultat d'un traitement cérébral très complexe. Depuis longtemps, les neurobiologistes travaillent pour comprendre l'anatomie et le fonctionnement du cerveau. Ils ont découvert une structure très complexe qui reste un mystère que les sciences tentent d'élucider. David Hubel a merveilleusement exprimé cette conquête [77] : le cerveau peut-il comprendre le cerveau ?

Avec l'apparition de machines de calcul de plus en plus puissantes et la mise en place de nouveaux moyens technologiques de traitement de l'information, de nombreux scientifiques se sont attaqués au problème de la vision d'un point de vue quantitatif. Il s'agit de créer des modèles capables de reproduire certaines fonctions du système visuel humain en se basant sur les résultats de la recherche dans le domaine de la vision biologique. Grâce aux efforts fournis par les chercheurs en robotique et intelligence artificielle, il est aujourd'hui possible de concevoir des systèmes de vision capables d'effectuer certaines tâches, comme la navigation d'un robot dans un environnement contrôlé ou encore la reconnaissance de certains objets dans une chaîne de production. Cependant, on est encore loin d'égaliser les performances de la vision humaine.

Un des objectifs de la vision artificielle est de percevoir la structure tridimensionnelle (3D) d'un environnement à partir d'une ou plusieurs images vidéo. La stéréovision constitue la technique la plus utilisée pour atteindre cet objectif. Il s'agit de reconstruire, partiellement ou entièrement, une scène 3D à partir de deux ou plusieurs images prises sous des angles légèrement différents. Le problème central de la stéréovision est celui de la mise en correspondance des indices visuels extraits des images stéréoscopiques. La re-

construction de la scène 3D est ensuite assurée par l'exploitation des décalages associés aux indices visuels appariés. Le présent chapitre décrit les principes fondamentaux de la stéréovision, avec une attention particulière au problème de la mise en correspondance.

## 1.2 Techniques pour la perception de l'environnement

La perception de l'environnement est une problématique étudiée depuis longtemps par les roboticiens. Avec l'émergence de l'idée de développer des systèmes d'aide à la conduite pour des véhicules routiers, cette problématique est devenue le centre d'intérêt de nombreux chercheurs. En effet, la perception de l'environnement d'un véhicule constitue le maillon fort de la chaîne de prise de décision dans un système d'aide à la conduite. L'assistance à l'utilisateur dépend fortement de cette étape clé où la précision, la robustesse et la rapidité doivent assurer une grande fiabilité.

Destinée à l'amélioration de la sécurité dans les systèmes de transports terrestres, cette nouvelle approche a pour objectif de créer une "voiture intelligente" capable de fournir des conseils à l'utilisateur sur la conduite à adopter, le prévenir de tout risque d'accident, voire prendre des initiatives pour éviter un accident imminent, grâce à l'utilisation de capteurs qui lui permettent de percevoir son environnement, de tester son propre état et celui de son conducteur. Plus sophistiqués encore, des systèmes de sécurité interactifs permettraient de connecter l'automobiliste à l'infrastructure routière ou à d'autres conducteurs pour des échanges d'informations portant sur l'état de la route, le trafic, les conditions de circulation, etc.

Dans ce cadre, plusieurs programmes et projets de recherche ont été lancés dans le monde entier pour répondre aux nombreuses questions liées au concept de la "voiture intelligente". On peut citer le programmes IVHS (Intelligent Vehicle and Highway System) aux Etats-Unis et VICS (Vehicle Information and Communication System) au Japon. En Europe, depuis PROMETHEUS (Program for an European Traffic with Highest Efficiency and Unprecedented Safety) et DRIVE (Dedicated Infrastructure for Vehicle Safety in Europe) qui étaient les premiers dans ce domaine, d'autres programmes ont été lancés : ROADSENSE (Road Awareness for Driving via a Strategy that Evaluates Nume-

rous Systems), RADARNET (Multifunctional Automotive Radar Network), CARSENSE (Sensing of Car Environment at Low Speed Driving), PREVENT (Preventive and Active Safety). En France, on peut citer les programmes PAROTO (Projet Anticollision Radar et Optronique pour l'Automobile), STATUE (Système Télématique destiné à l'Accroissement de la sûreté des Transports guidés Urbains et à l'Emergence de nouvelles aides à l'exploitation), ARCOS (Action de Recherche pour une Conduite Sécurisée) et MobiVIP (Véhicules Individuels Public pour la Mobilité en centre ville), soutenus par le programme national PREDIT (Programme pour la Recherche, le Développement et l'Innovation dans les Transports terrestres).

L'objectif de base de la perception de l'environnement d'un véhicule est de détecter, localiser et/ou suivre le mouvement des objets fixes ou mobiles présents autour du véhicule. De nombreuses techniques ont été développées pour parvenir à réaliser une ou plusieurs de ces fonctions. Elles relèvent, selon la technologie utilisée, d'une des deux approches suivantes : approche active ou passive.

### 1.2.1 Approche active

L'approche active est basée sur l'utilisation de systèmes dotés d'un émetteur et d'un récepteur. L'émetteur envoie vers la scène un faisceau d'ondes radio (radar), d'ondes sonores (sonar) ou de lumière (lidar), dont les éventuelles rétro-diffusions vers le récepteur permettent de localiser les objets présents dans la scène. En mesurant le temps aller-retour du faisceau émis, on peut alors connaître la distance des objets par rapport au capteur.

Utilisés pour la détection d'obstacles, les systèmes actifs sont caractérisés par la puissance et l'angle d'ouverture de l'émetteur, la sensibilité et la sélectivité du récepteur. Les performances de ces techniques actives peuvent être réduites par certains problèmes liés, d'une part, au milieu de propagation qui absorbe et/ou disperse les faisceaux aller et retour, et d'autre part, à la surface réfléchissante de la cible qui peut absorber le faisceau, le diffuser ou le réfléchir dans d'autres directions.

De nombreux systèmes actifs de détection d'obstacles sont proposés dans la littérature : radar [39, 38, 157], sonar [147, 95], lidar [167, 64, 43].

### 1.2.2 Approche passive

Ne nécessitant l'émission d'aucun signal, les systèmes passifs ne sont constitués que d'une ou plusieurs caméras. Les avantages par rapport aux systèmes actifs sont un affranchissement des problèmes d'inférences capteurs/capteurs ou capteurs/environnement, une meilleure résolution et une exploration rapide de la scène observée. L'approche passive présente néanmoins des inconvénients liés généralement à une segmentation difficile des images à cause du bruit ou d'un faible contraste entre les objets de la scène. Pour les scènes d'extérieur, la route par exemple, une mauvaise visibilité, la nuit ou dans certaines conditions climatiques (pluie, brouillard, neige), rend souvent les images inexploitable et compromet ainsi la qualité de la perception.

Le principe de l'approche passive consiste à traiter et analyser les images issues des caméras. Si l'aspect algorithmique est un axe de recherche fondamental dans cette approche, les travaux portent également sur l'implantation temps-réel des algorithmes sur des architectures matérielles dédiées. Plusieurs techniques à base de traitement et analyse d'images ont été développées pour la perception de l'environnement d'un véhicule [41, 154, 171, 94, 3, 152, 9, 13, 93]. D'autres techniques combinent les systèmes actifs et passifs [84, 44, 108, 55, 24].

Une des techniques passives les plus connues est la vision stéréoscopique. La suite de ce chapitre est consacrée à la présentation de cette technique.

## 1.3 Processus de stéréovision

On désigne par stéréovision, vision stéréoscopique ou stéréoscopie, l'ensemble de tous les procédés qui permettent de déterminer la structure tridimensionnelle d'une scène à partir de plusieurs images vidéo de celle-ci, prises sous des angles de vue différents. D'ailleurs, le mot "stéréoscopie" qui dérive des mots grec "stereos", solide, et "skopein", voir, signifie "voir en relief".

Le système stéréoscopique le plus courant et le plus simple n'utilise que deux images. Il s'agit alors de la stéréovision binoculaire. Cependant, certains auteurs préconisent l'utilisation de trois caméras pour lever certaines ambiguïtés en cas d'occultations [4]. On

parle alors de la stéréovision trinoculaire.

Comme chez l'Homme, la perception du relief par stéréovision binoculaire est assurée principalement par l'exploitation du décalage existant entre les deux images stéréoscopiques. Ce décalage, appelé disparité stéréoscopique, ne peut être évalué qu'en établissant des correspondances entre les deux images.

### 1.3.1 Décomposition du processus de stéréovision

Le problème clé de la stéréovision est de trouver dans les deux vues, les images des mêmes entités physiques de la scène observée. Sans contraintes particulières, une mise en correspondance point à point des images n'est pas une solution idéale. En effet, pour des images de taille courante de 512x512 pixels, on dénombre  $512^4$ , soit 68 milliards de couples de points à tester. De plus, les pixels des images ne sont pas tous significatifs, surtout pour des applications telles que la détection d'obstacles. On procède alors à l'appariement de primitives-image extraites des images (points remarquables, contours, régions, etc.). Constituant une description de la scène perçue, les primitives-objet sont reconstruites en utilisant des calculs de triangulation géométrique sur les paires de primitives-image homologues. Ces calculs nécessitent la définition de systèmes de coordonnées (repères) liés aux images et à la scène observée, et donc la connaissance de la géométrie du capteur stéréoscopique.

Le processus de stéréovision peut se décomposer alors en quatre étapes :

1. Choix d'un modèle géométrique du capteur stéréoscopique.
2. Extraction des primitives-image.
3. Mise en correspondance des primitives-image.
4. Triangulation géométrique et reconstruction tridimensionnelle de la scène à partir des primitives-objet.

### 1.3.2 Principes généraux

Les stratégies utilisées pour aborder le problème de mise en correspondance stéréoscopique sont nombreuses. Elles dépendent généralement des réponses qu'elles apportent

aux questions suivantes :

- Quelles primitives choisir pour l'appariement ? Ce choix conditionne le problème de l'appariement. En effet, certains types de primitives sont plus aisés à appairer que d'autres. De plus, ce choix détermine la représentation finale de la scène observée. Les différentes primitives utilisées en stéréovision sont présentées dans la section 1.5.
- Quels critères utiliser ? Afin de mesurer la qualité des appariements, un ou plusieurs critères de ressemblance doivent être optimisés. La section 1.6 présente différents critères utilisés en stéréovision.
- Quelles contraintes utiliser ? Afin de lever d'éventuelles ambiguïtés dans les appariements, un certain nombre de contraintes doivent être satisfaites. Les différentes contraintes employées dans la littérature sont présentées dans la section 1.7.
- Quelles méthodologies adopter ? C'est un autre aspect différenciant les algorithmes proposés pour résoudre le problème de mise en correspondance. En effet, il n'existe pas actuellement de méthode suffisamment générale qui puisse s'appliquer à une paire d'images stéréoscopiques, indépendamment de leur contenu et de leur type. La section 1.8 présente les différentes techniques de mise en correspondance utilisées en stéréovision.

### 1.3.3 Aspects fondamentaux de l'appariement

#### 1.3.3.1 Fonction d'appariement

Soient deux listes  $G$  et  $D$  de primitives gauche et droite d'un couple d'images stéréoscopiques. L'appariement entre ces deux listes peut être défini comme une fonction de  $G$  vers  $D$  (ou inversement). La qualité de cette fonction d'appariement se mesure par :

1. Le nombre de primitives appariées, qui doit être aussi élevé que possible.
2. Le nombre de faux appariements, qui doit être aussi faible que possible.

La recherche des appariements n'est pas la recherche d'une bijection stricte entre les listes  $G$  et  $D$ . Seules certaines primitives, de chaque image, seront appariées. En effet, un élément d'une image peut ne pas être vu dans l'autre image et vice-versa, soit parce qu'il

est caché par un autre élément, soit parce qu'il se trouve exclusivement dans le champ de vue d'une seule caméra, soit enfin à cause de phénomènes physiques tels que les ombres et les reflets qui dépendent des positions relatives de la source lumineuse, de l'objet observé et de la caméra.

### 1.3.3.2 Complexité du problème de l'appariement

Le problème de mise en correspondance est de nature combinatoire : si le nombre de primitives de l'image gauche (respectivement de l'image droite) est  $N_g$  (respectivement  $N_d$ ), le nombre de couples possibles est alors  $N_g \times N_d$ . Parmi ces couples, seules  $N$  paires seront retenues avec, si l'on prend en considération les remarques du paragraphe précédent,  $N \leq \min(N_g, N_d)$ . Il est donc fondamental de réduire la combinatoire, d'une part, pour diminuer le nombre d'erreurs d'appariement possibles, et d'autre part, pour limiter le temps des traitements.

### 1.3.3.3 Dilemme précision – appariement

Après la phase de mise en correspondance, les couples de primitives appariées permettent, par triangulation, de calculer les positions des primitives-objet correspondantes dans la scène réelle. La précision de la localisation est accrue si les points de vue sont nettement différents, c'est à dire lorsque la distance séparant les centres optiques des deux caméras est importante. Mais plus les caméras sont éloignées l'une de l'autre, plus les images stéréoscopiques sont dissemblables. L'apparence des surfaces et des contours des objets diffèrent alors sensiblement et le champ de vision commun est réduit. D'importantes déformations géométriques et occultations sont également plus probables. Par conséquent, la tâche de mise en correspondance devient plus difficile à réaliser. Il faut donc trouver, selon l'application envisagée, un compromis entre précision et facilité d'appariement.

## 1.4 Aspects géométriques de la stéréovision

### 1.4.1 Géométrie du capteur stéréoscopique

La géométrie du système de prise de vue est un élément important dans un processus de stéréovision. Pour que deux caméras constituent un capteur stéréoscopique, il faut que celles-ci observent la même scène, c'est à dire qu'elles soient disposées de telle sorte que leurs axes optiques convergent vers la scène observée (cf. figure 1.1). En utilisant le modèle sténopé, un point  $P$  de la scène donne naissance à deux projections  $p_g$  et  $p_d$  sur les images gauche et droite ( $p_g$  et  $p_d$  sont appelés des points homologues). On peut ainsi déterminer les équations de la droite  $O_g p_g$  passant par le centre optique  $O_g$  de la caméra gauche et le point-image  $p_g$ , et de la droite  $O_d p_d$  passant par le centre optique  $O_d$  de la caméra droite et le point-image  $p_d$ . L'intersection de ces deux droites donne, par triangulation géométrique, le point  $P$  de la scène réelle. Afin de pouvoir calculer cette intersection et donc de déterminer la position de  $P$ , il faut exprimer les deux équations des droites  $O_g p_g$  et  $O_d p_d$  dans un même repère. On choisit en général un repère associé à la scène, appelé repère de référence. Il est donc nécessaire de connaître les relations permettant de passer du point  $P$  de l'espace 3D, exprimé dans le repère de référence, vers les points-image  $p_g$  et  $p_d$ , exprimés dans les repères 2D liés aux images gauche et droite. La détermination de ces relations s'effectue lors de la phase de calibrage du système de prise de vue stéréoscopique.

### 1.4.2 Calibrage

Le calibrage d'un capteur stéréoscopique consiste à déterminer les transformations géométriques permettant de passer d'un point de l'espace 3D, exprimé dans un repère de référence lié à la scène observée, vers ses projections sur les images gauche et droite, exprimées dans les repères 2D liés aux images. Pour cela, il faut disposer de paramètres intrinsèques et de paramètres extrinsèques. Les paramètres intrinsèques décrivent les caractéristiques propres à chaque caméra, indépendamment de leur association stéréoscopique. Il s'agit de la distance focale des objectifs des caméras, de la taille et de la résolution des

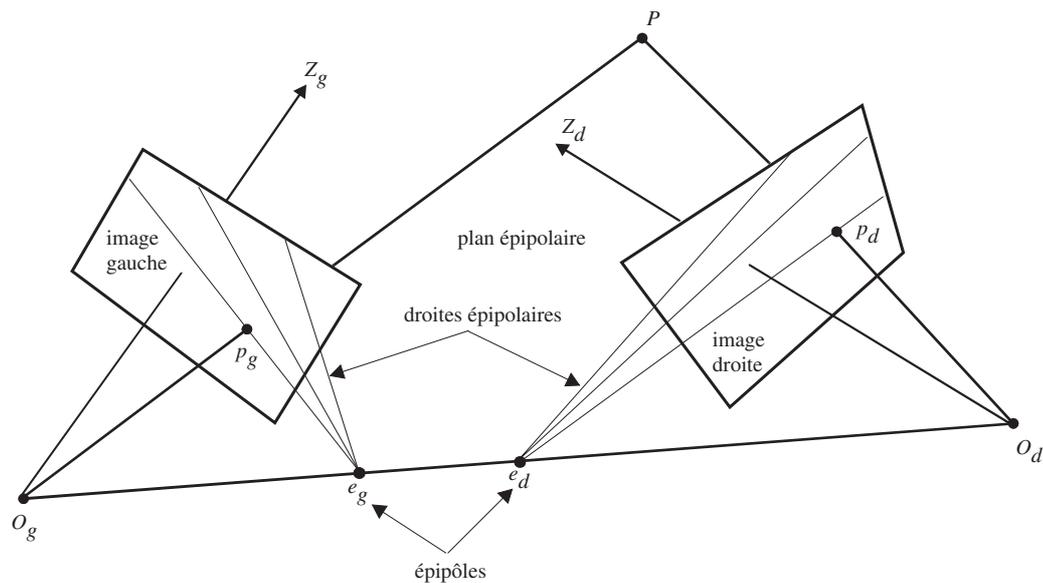


FIG. 1.1 – Configuration générale d'un capteur stéréoscopique.

capteurs CCD (Charge Coupled Device). Les paramètres extrinsèques définissent l'association stéréoscopique des caméras et expriment la position et l'orientation de chacune d'elles par rapport au repère de référence.

La transformation associée à chaque caméra est composée de deux matrices. La première, déduite des paramètres extrinsèques, permet le passage du repère de référence au repère lié à la caméra. La deuxième matrice, calculée à partir des paramètres intrinsèques, définit le passage d'un point de l'espace 3D, exprimé dans le repère lié à la caméra, vers sa projection sur l'image, exprimée dans le repère 2D lié à l'image.

Pour déterminer ces transformations, plusieurs techniques de calibrage ont été proposées dans la littérature [159, 67, 179].

### 1.4.3 Droites épipolaires

La notion de droite épipolaire joue un rôle très important en stéréovision. Elle permet de réduire l'espace des solutions lors de la recherche des appariements. Une droite épipolaire décrit le lieu des points d'une image pouvant correspondre à un même point de l'autre image (cf. figure 1.1). Il y a donc, pour chaque point de l'image gauche, une droite épipolaire dans l'image droite et, réciproquement, pour chaque point de l'image droite, il

Il y a une droite épipolaire dans l'image gauche. De plus, deux points d'une droite épipolaire d'une image ont leurs correspondants, s'ils existent, sur une droite épipolaire de l'autre image. Ces deux droites sont appelées droites épipolaires conjuguées. Dans chaque image, les droites épipolaires forment un faisceau de droites qui convergent vers un même point, appelé épipôle (cf. figure 1.1). En fait, l'épipôle droit  $e_d$ , par exemple, n'est que la projection du centre optique  $O_g$  de la caméra gauche sur l'image droite. On peut également définir l'épipôle gauche  $e_g$  de la même manière.

#### 1.4.4 Configuration idéale

Dans le cas d'une configuration quelconque des caméras (cf. figure 1.1), il est nécessaire de calculer pour chaque point d'une image l'équation de la droite épipolaire correspondante dans l'autre image. Cependant, il existe une configuration particulièrement intéressante permettant d'avoir des droites épipolaires parallèles et confondues avec les lignes des images. Cette configuration, où les droites épipolaires conjuguées sont confondues, est obtenue quand les axes des deux caméras sont parallèles et quand la droite  $O_gO_d$  reliant les centres optiques est confondue avec les axes horizontaux des deux caméras (cf. figure 1.2). Dans ce cas, les deux épipôles se trouvent rejetés à l'infini.

Les droites épipolaires jouent un rôle fondamental en vision stéréoscopique. En effet, lorsqu'on cherche pour un point d'une image un correspondant dans l'autre image, on peut limiter cette recherche le long de la droite épipolaire correspondante. Il est donc important d'avoir les expressions mathématiques les plus simples pour ces droites épipolaires. Comme on ne peut pas toujours disposer les deux caméras pour obtenir la configuration idéale de la figure 1.2, de nombreux auteurs [4, 72] préfèrent partir d'une configuration quelconque (cf. figure 1.1) et se ramener de façon analytique à la configuration idéale (cf. figure 1.2) en effectuant une rectification épipolaire. Celle-ci consiste à appliquer aux images une transformation linéaire en coordonnées projectives de façon à obtenir une paire d'images stéréoscopiques coplanaires et parallèles à la droite passant par les centres de projection des deux caméras. Il faut toutefois noter que cette opération implique un accroissement du temps de calcul [96]. Plusieurs techniques de rectification d'images stéréoscopiques sont proposées dans la littérature [52, 65].

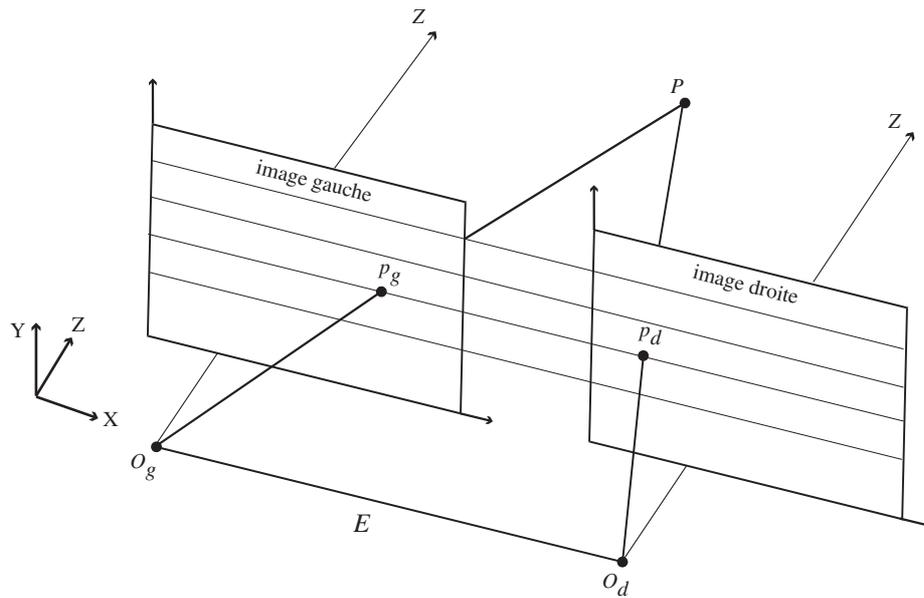


FIG. 1.2 – Configuration particulière du système de prise de vue.

### 1.4.5 Reconstruction tridimensionnelle

A l'issue de la phase de mise en correspondance, les couples de primitives-image appariées sont utilisés pour calculer les positions 3D des primitives-objet. Ce calcul, appelé triangulation géométrique, est possible grâce aux transformations qui sont établies lors de la phase de calibrage.

Certaines applications nécessitent une reconstruction surfacique de la scène observée par le capteur stéréoscopique. Il s'agit de fournir une représentation synthétique et formelle de la scène. Pour cela, on utilise les positions tridimensionnelles des primitives-objet calculées par triangulation. Il faut noter que la reconstruction 3D dépend du type de primitive utilisé lors de la phase de mise en correspondance. Le type de primitive choisi détermine alors la représentation finale de la scène. De plus, la présence d'un grand nombre de primitives densifie la reconstruction 3D mais, en même temps, augmente les calculs nécessaires à la reconstruction. En revanche, lorsque les primitives ne sont pas nombreuses, les calculs sont réduits mais la reconstruction est moins riche.

La reconstruction 3D a pour objectif de reconstituer les différents objets présents dans la scène. Sans relations entre les primitives-objet reconstruites et les objets physiques, la reconstruction est impossible. D'où la nécessité de disposer d'un modèle de la

scène, c'est à dire d'un ensemble de connaissances à priori sur les objets à reconstruire.

## 1.5 Les primitives

Avant d'envisager l'étape de mise en correspondance, on cherche tout d'abord à associer une représentation symbolique aux images stéréoscoiques. Il s'agit d'extraire de chaque image des indices visuels, appelés primitives, caractérisant les objets présents dans l'image. Mettre en correspondance deux images stéréoscopiques revient alors à mettre en correspondance les primitives extraites de chacune d'elles.

Les primitives les plus simples ne nécessitant aucun pré-traitement des images sont les pixels. Cependant, un tel choix présente les inconvénients suivants :

1. L'information associée à un pixel, le niveau de gris ou la couleur, n'est pas suffisamment discriminante pour lever les ambiguïtés lors de l'appariement. Il peut donc y avoir de nombreux appariements erronés.
2. L'image contient en général un nombre très important de pixels. Par conséquent, leur appariement conduit à une combinatoire importante et donc à un accroissement considérable des temps de calcul. En fonction de l'application envisagée, cette limitation peut être levée par l'utilisation d'algorithmes de mise en correspondance offrant la possibilité d'être implantés sur des architectures dédiées.

### 1.5.1 Les propriétés requises

Afin d'optimiser le processus de stéréovision en termes de qualité d'appariement, de temps de traitement et de bonne localisation, les primitives à apparier doivent vérifier certaines propriétés [4]. Elles doivent être :

- **Discriminantes** : Les primitives doivent avoir des attributs qui permettent de les discriminer afin de lever toute ambiguïté lors de la mise en correspondance.
- **Denses** : Les primitives doivent être suffisamment nombreuses pour représenter tous les objets de la scène.
- **Compactes** : L'image doit être représentée d'une manière concise afin de réduire

la complexité algorithmique de la procédure de mise en correspondance.

- **Précises** : La qualité de localisation des objets physiques dépend de la position des primitives qui doivent donc être extraites avec précision.
- **Robustes** : Les primitives doivent être peu sensibles aux petites variations d'intensité dans l'image dues au bruit d'acquisition, de numérisation, etc.
- **Intrinsèques** : Les primitives doivent correspondre à la projection, dans l'image, des objets physiques.

## 1.5.2 Types de primitives

Le choix des primitives est fondamental en stéréovision. Il est conditionné d'une part par le type d'images à traiter, mais également par l'information que l'on désire en extraire afin d'obtenir la représentation finale de la scène 3D. Après la mise en correspondance, la phase de modélisation peut être immédiate si le résultat de la triangulation des primitives appariées donne directement les modèles tridimensionnels utiles. Toutefois, des étapes intermédiaires de mise en forme des données sont parfois nécessaires suivant le niveau de représentativité des primitives-objet reconstruites. Cette section présente les principaux types de primitives utilisés en stéréovision.

### 1.5.2.1 Les points d'intérêt

Comme il a été mentionné au début de la section 1.5, l'un des inconvénients liés à l'utilisation du pixel comme primitive est l'augmentation de la combinatoire de la procédure de mise en correspondance. Cependant, certains auteurs ont proposé de limiter le traitement à un certain nombre de pixels, appelés points d'intérêt, possédant des propriétés et des caractéristiques les distinguant de leurs voisins. Les attributs associés à ces points sont, comme pour les pixels, de nature photométrique (intensité lumineuse, contraste, etc.). Les primitives de ce type sont les premières à avoir donné des résultats satisfaisants en stéréovision. Parmi les plus connues, on peut citer les points de Moravec [113] et les points de forte courbure [122].

### 1.5.2.2 Les contours

En raison de leur représentativité, les primitives de type contour sont les plus utilisées en stéréovision. Elles représentent les frontières des objets et les variations spatiales des propriétés géométriques et physiques de leurs surfaces (orientation, matière, texture, couleur, etc.). De cette façon, l'extraction des contours permet de mettre facilement en évidence les éléments significatifs d'une image.

Il faut cependant remarquer que la notion de contour est double. Elle englobe, d'une part, l'ensemble des points constituant le contour, et d'autre part, les modèles de contours des objets obtenus à partir de ces points.

Les points de contour correspondent aux variations brusques de la fonction niveau de gris de l'image. Ils s'obtiennent généralement en utilisant des techniques de filtrage (recherche des extrema locaux du gradient, recherche des passages par zéro du Laplacien, etc.) et de seuillage [119, 22, 40, 166, 28]. En plus de son niveau de gris, un point de contour peut être caractérisé également par son gradient qui donne une information sur la direction du contour de l'objet en ce point.

Cependant, l'image des points de contour est parfois inexploitable. En effet, des points indésirables viennent s'ajouter aux points de contour effectifs. D'autre part, certains points de contour ne sont pas détectés. De plus, le nombre de points de contour obtenus est généralement très important, ce qui augmente la combinatoire de la mise en correspondance. Pour pallier ces inconvénients, certains auteurs préfèrent chaîner les points de contour et modéliser les contours par approximation sous forme de segments, de polygones, d'arcs de cercles, etc. [110, 70, 4, 25, 56, 28]. Ainsi, les contours permettent de fournir une représentation relativement synthétique de l'information contenue dans l'image. Ils sont décrits par des attributs de type photométrique, mais aussi de type géométrique tels que la longueur, l'orientation, la courbure, etc. On peut leur associer également d'autres attributs plus globaux décrivant les relations structurelles qui les relient [151]. Toutefois, les contours présentent un inconvénient lié aux effets de perspective. En effet, les images d'un même contour vu par plusieurs caméras n'ont généralement pas la même forme. En pratique, l'utilisation des contours est surtout intéressante avec des

objets polyédriques [72].

### 1.5.2.3 Les régions

Dans une image, une région est un ensemble de pixels adjacents, satisfaisant un prédicat qui indique habituellement une uniformité de certaines caractéristiques, telles que le niveau de gris, la couleur, la texture, etc. Ce type de primitive peut être caractérisé par un ensemble d'attributs de nature photométrique (niveau de gris moyen, texture, etc.), géométrique (surface, forme) ou morphologiques (compacité, moments d'inertie, coordonnées du centre d'inertie). Pour les images couleurs, on peut ajouter d'autres informations de type colorimétrique. L'utilisation des régions en stéréovision est avantageuse, car elle permet de diminuer les possibilités d'appariement et donc de réduire la combinatoire de la procédure de mise en correspondance. En contrepartie, comme pour les contours, leur appariement est souvent rendu difficile par le phénomène de distorsion affectant des images saisies sous des angles différents.

Les surfaces des objets n'ont pas toujours un niveau de gris ou une couleur uniforme, mais peuvent présenter un certain motif ou texture. La segmentation d'images utilisant l'analyse de texture [20] est intéressante lorsqu'on choisit les régions comme primitives. En effet, quand il s'agit d'images fortement texturées, il est plus facile de les décrire par des régions de textures homogènes que par des contours qui, dans ce cas, seront très nombreux. En revanche, l'analyse de texture est souvent très lente en raison de l'utilisation d'outils très coûteux en temps de calculs (transformée de Fourier, matrices de co-occurrence). Plusieurs techniques de segmentation d'images en régions sont proposées dans la littérature [10, 28].

Notons au passage que l'utilisation de la couleur est peu fréquente, car elle nécessite du matériel coûteux et implique de travailler sur trois plans-image, ce qui rend les traitements plus lourds. Mais l'information couleur peut être d'un grand intérêt lors de l'appariement de primitives. L'analyse d'images en couleur est devenue, depuis quelques années, l'un des domaines sur lequel de nombreux chercheurs se sont penchés, mettant au point de nombreux outils pouvant servir en stéréovision [97, 160, 162, 49].

#### 1.5.2.4 Les contours actifs

Initialement introduits par Kass [83], les modèles de contours actifs, ou "snakes" [173, 101, 30, 138, 131], permettent, par rapport aux primitives précédentes, une représentation plus globale des objets. L'idée de base consiste à approcher le contour d'un objet par une courbe continue, fermée ou non. Celle-ci est constituée d'un ensemble de points de contrôle, appelés "snaxels".

Le processus de recherche du contour actif s'effectue en deux étapes. La première étape concerne l'initialisation du contour actif au voisinage du contour à extraire. La deuxième étape consiste à faire évoluer le contour actif jusqu'à ce qu'il coïncide avec le contour recherché. L'évolution du contour actif est régie par la minimisation de son énergie, composée de deux termes, l'un correspondant à une énergie externe, l'autre à une énergie interne. L'énergie externe permet d'attirer le contour actif vers le contour de l'objet. L'énergie interne est destinée à donner une certaine régularité au contour actif.

Dans sa thèse, Selsis [146] a utilisé les contours actifs pour suivre et localiser par stéréovision des objets en mouvement. En faisant l'hypothèse que tous les objets en mouvement pénètrent dans les champs des caméras par leurs périphéries, elle a proposé une initialisation automatique en utilisant un contour actif placé à la périphérie des images. Ainsi, tout objet pénétrant dans le champ de chaque caméra déforme ce contour actif périphérique jusqu'à sa fermeture autour de l'objet. Une procédure de scission permet alors d'initialiser un contour actif propre à cet objet, afin de le suivre durant sa traversée du champ de vision. Le contour actif périphérique retrouve, quant à lui, sa position initiale, afin de modéliser les nouveaux objets pénétrant dans la scène. Afin de localiser les objets en mouvement, Selsis a utilisé les contours actifs comme primitives lors de la phase de mise en correspondance stéréoscopique. Trois attributs sont retenus pour caractériser ces primitives : la forme, la surface et les coordonnées du barycentre.

## 1.6 Les critères de ressemblance

La ressemblance est une notion utilisée par la plupart des algorithmes de mise en correspondance. L'hypothèse sous-jacente à la ressemblance est que les attributs attachés

à deux primitives qui se correspondent doivent être semblables. L'idée de base consiste à apparier une primitive en la comparant à ses correspondants possibles, de telle sorte que le couple qui se ressemble le plus soit retenu comme un couple de primitives homologues. Il faut donc, dans un premier temps, établir un critère permettant de quantifier la ressemblance. Ce critère doit tendre vers une valeur limite pour une ressemblance parfaite. En fait, cette valeur n'est jamais atteinte car deux primitives homologues ne sont jamais parfaitement identiques. Etant donnée une primitive, la recherche de son homologue consiste donc à sélectionner parmi les appariements possibles celui qui optimise le critère de ressemblance. Selon le type de primitives utilisé, les critères de ressemblance peuvent être de nature photométrique ou géométrique.

- **Critères photométriques** : Ce sont les critères les plus utilisés lorsqu'il s'agit des primitives de type point. Le principe consiste à examiner les niveaux de gris sur des fenêtres de même taille, centrées sur les points que l'on cherche à apparier. La ressemblance de deux points peut alors se mesurer au moyen d'un coefficient de corrélation du contenu de ces deux fenêtres. Des exemples de critères de corrélation sont présentés dans [46, 72]. Il faut noter que la corrélation présente l'inconvénient de travailler nécessairement sur des fenêtres de tailles identiques. Or, dans le cas des images stéréoscopiques, les effets du phénomène de distorsion projective conduisent deux primitives homologues à être de tailles et de formes différentes. Pour pallier à cette limitation, Ohta [124] propose d'utiliser un critère de pseudo-corrélation qui prend en compte ces effets de perspective.
- **Critères géométriques** : Ces critères s'appliquent généralement aux primitives de haut niveau, de type contour ou région par exemple, caractérisées par des attributs géométriques tels que la longueur des segments, l'angle entre les segments, la surface ou le périmètre des régions, etc. Il est donc possible de déterminer différents critères prenant en compte ces caractéristiques géométriques. Ces critères permettent de déterminer les couples de primitives candidats à l'appariement en utilisant des seuils de tolérance. Ces derniers peuvent être déduits de la géométrie épipolaire du capteur stéréoscopique [45].

## 1.7 Les contraintes

L'utilisation des critères de ressemblance n'est en général pas suffisante pour établir des appariements sans ambiguïté. Une primitive de l'image gauche peut en effet ressembler à plusieurs primitives de l'image droite et réciproquement. Afin de lever les ambiguïtés d'appariement, on exploite un certain nombre de contraintes de type locale et globale.

### 1.7.1 Les contraintes locales

Les contraintes locales sont basées sur la comparaison des attributs des primitives. Elles servent à identifier les couples de primitives candidats à l'appariement en ne retenant que ceux qui présentent des attributs semblables. Ces contraintes sont utilisées généralement pour réduire l'espace de recherche des primitives homologues. Parmi les contraintes locales les plus connues, on peut citer les contraintes épipolaire, d'orientation et de disparité maximale et/ou minimale.

#### 1.7.1.1 La contrainte épipolaire

La contrainte épipolaire est la seule contrainte géométrique intrinsèque au capteur stéréoscopique. Elle est très souvent utilisée car elle simplifie la recherche des correspondants lors de l'appariement des primitives. Comme il a été mentionné dans la section 1.4.3, dans le cas des primitives de type point, les droites épipolaires conjuguées se caractérisent par le fait que le correspondant d'un point de l'une se trouve, s'il existe, exclusivement sur l'autre. D'où l'importance de cette contrainte qui permet de limiter la recherche des homologues à une droite au lieu de la totalité de l'image. Cette contrainte peut être appliquée à d'autres primitives telles que les segments [146] et les régions [10].

#### 1.7.1.2 La contrainte d'orientation

La contrainte d'orientation impose que les primitives candidates à l'appariement présentent des orientations semblables. Pour les segments de droite, les orientations sont définies par l'angle entre le vecteur directeur associé au segment et un des axes du repère de l'image. Dans le cas des primitives de type point, le vecteur directeur est celui de

la tangente au contour au point considéré. Lorsqu'il s'agit d'apparier des régions, cette contrainte peut être utilisée en considérant les directions de leurs axes principaux et secondaires. Cependant, les déformations des régions dues aux effets de perspective peuvent rendre cette contrainte inapplicable. La compatibilité d'orientation de deux primitives gauche et droite se mesure par une simple différence de leurs orientations respectives. L'appariement de deux primitives est rejeté si la différence de leurs orientations dépasse une valeur limite [72].

### 1.7.1.3 La contrainte de disparité maximale et/ou minimale

La contrainte de disparité est utilisée lorsque l'on connaît les limites de la zone de la scène utile à l'application envisagée. Cette zone est généralement définie par une distance minimale et une distance maximale par rapport au dispositif de prise de vue. La distance étant inversement proportionnelle à la disparité, la procédure de mise en correspondance ne considère alors que les couples de primitives ayant des disparités comprises entre une valeur minimale et une valeur maximale [90].

## 1.7.2 Les contraintes globales

Les contraintes globales s'appliquent à l'ensemble des appariements possibles obtenus après l'exploitation des contraintes locales. Elles servent à valider la compatibilité entre les couples de primitives en levant les ambiguïtés d'appariement. Les contraintes globales les plus utilisées en stéréovision sont celles d'unicité, d'ordre, de continuité de la disparité et de continuité figurale.

### 1.7.2.1 La contrainte d'unicité

Comme il a été mentionné dans la section 1.3.3.1, la fonction d'appariement n'est pas une bijection stricte entre les ensembles de primitives de l'image gauche et de l'image droite. En d'autres termes, certaines primitives d'une image n'auront pas de correspondants dans l'autre image et vice-versa. Tenant compte de cet aspect fondamental de la stéréovision, la contrainte d'unicité exprime le fait qu'une primitive d'une image a au plus

un correspondant dans l'autre image, et inversement.

### 1.7.2.2 La contrainte d'ordre

La contrainte d'ordre impose que les primitives appariées le long de droites épipolaires conjuguées soient dans le même ordre. Dans le cas des primitives de type point, deux couples  $(P_g^1, P_d^1)$  et  $(P_g^2, P_d^2)$  respectent la contrainte d'ordre si et seulement si la position ("à gauche de" ou "à droite de") le long de la droite épipolaire gauche de  $P_g^1$  par rapport à  $P_g^2$  est la même que celle de  $P_d^1$  par rapport à  $P_d^2$  le long de la droite épipolaire droite. La figure 1.3 illustre le cas où la contrainte d'ordre est respectée alors que la figure 1.4 représente un cas où elle n'est pas respectée.

Il est important de noter que cette contrainte n'est pas toujours valable, surtout en présence d'objets ponctuels (cf. Figure 1.5). Dans une scène naturelle, ce type de situation ne peut avoir lieu que de façon exceptionnelle puisque les objets ponctuels y sont plutôt rares. En effet, les points utilisés pour l'appariement sont associés aux objets présents dans la scène. On doit donc tenir compte de la taille de ces objets. La figure 1.6 illustre une situation identique à celle de la figure 1.5, mais en considérant des objets non ponctuels. L'objet situé à l'avant plan occulte celui qui se trouve à l'arrière plan. Ce dernier n'a donc plus d'image sur le capteur droit. Dans ce cas, le problème de l'ordre ne se pose plus. Il faut préciser que l'on exclut la présence d'objets transparents qui rendraient ce raisonnement inexact. Même avec des objets non ponctuels, on peut se retrouver dans des situations telles que celle illustrée sur la figure 1.7 où la contrainte d'ordre n'est pas applicable. Ce type de situation devient de plus en plus rare lorsque la distance entre le capteur stéréoscopique et la scène observée devient de plus en plus importante, comme dans le cas des images aériennes.

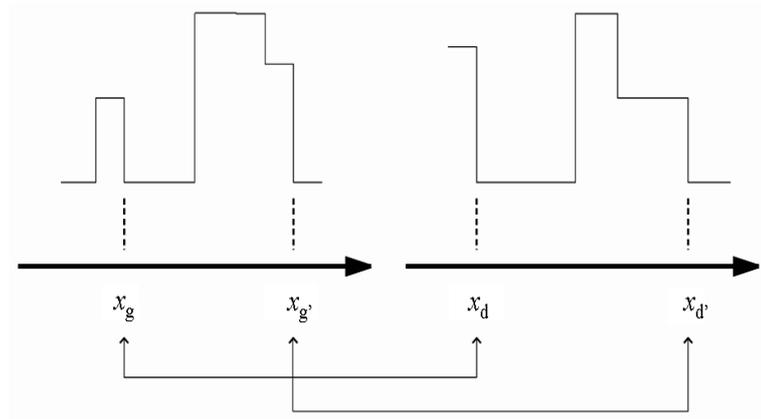


FIG. 1.3 – L'ordre est respecté.

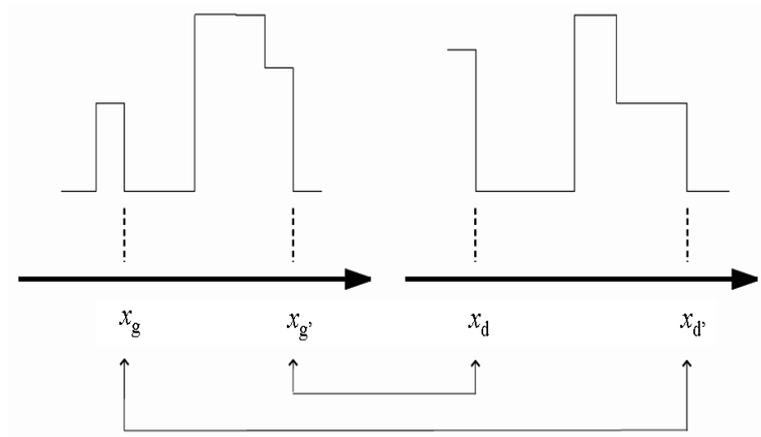


FIG. 1.4 – L'ordre n'est pas respecté.

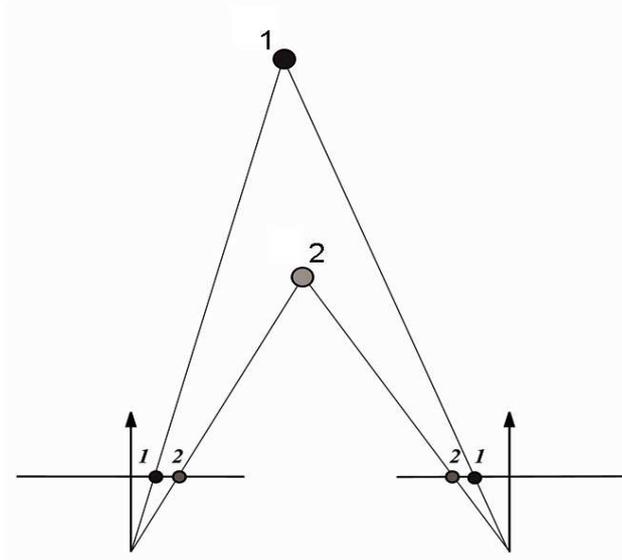


FIG. 1.5 – Cas d’objets ponctuels : la contrainte d’ordre n’est plus valable.

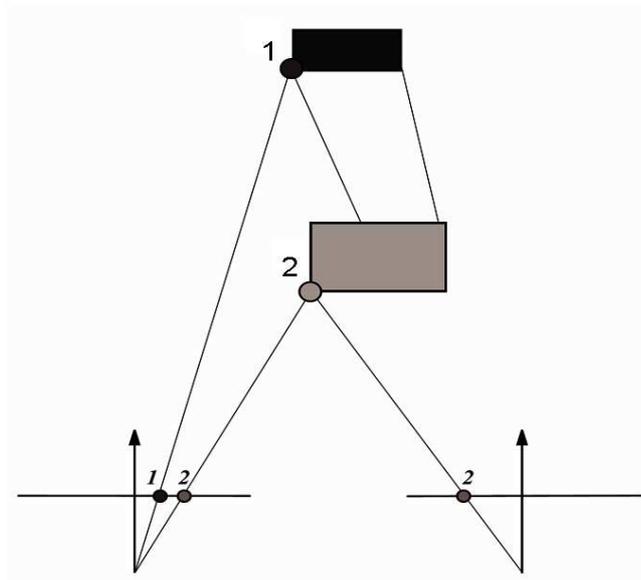


FIG. 1.6 – Cas d’objets non ponctuels : la validité de la contrainte d’ordre est assurée.

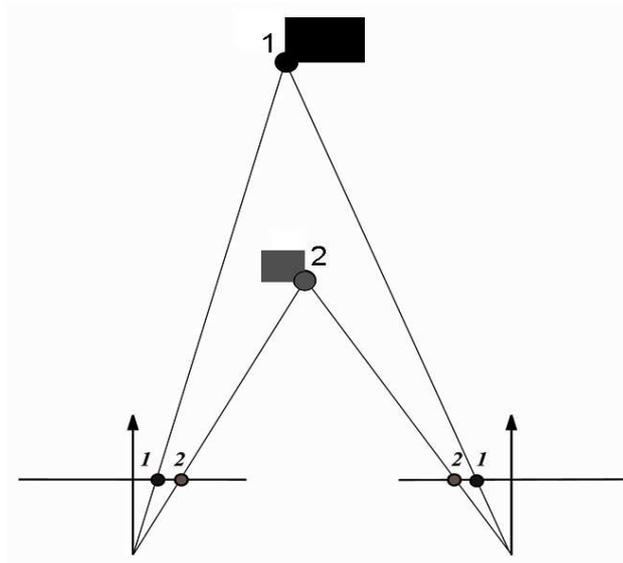


FIG. 1.7 – Cas d’objets non ponctuels : la contrainte d’ordre n’est pas applicable.

### 1.7.2.3 La contrainte de continuité de la disparité

L’application de la contrainte de continuité de la disparité suppose que la scène est formée d’objets physiques dont on peut approcher la structure spatiale par des primitives-objet, correspondant aux primitives visuelles des images [4]. Si on suppose que la surface des objets est suffisamment régulière, et que la densité des primitives-objet est suffisamment grande, les objets peuvent être représentés par des sous-ensembles de primitives-objet voisines dans l’espace, que l’on peut parcourir de manière régulière, presque continue. En d’autres termes, la variation de la profondeur, par rapport au dispositif de prise de vue, le long des primitives-objet voisines varie d’une manière presque continue. Ainsi, la disparité, liée directement à la profondeur, doit varier continûment le long des primitives-image voisines issues de primitives-objets voisines, c’est à dire que deux couples de primitives-image issues de deux primitives-objets voisines doivent avoir des disparités voisines. Certains auteurs ont exprimé cette contrainte à l’aide du gradient de disparité [4, 72], d’autres ont établi une caractérisation analytique grâce à la notion du gradient de disparité limite [135].

#### 1.7.2.4 La contrainte de continuité figurale

La contrainte de continuité figurale est généralement utilisée lorsque les primitives à mettre en correspondance se trouvent le long des contours des objets [72]. Considérons le cas où les droites épipolaires sont confondues avec les lignes des images stéréoscopiques. Si on suppose que la scène est constituée d'objets continus par morceaux, l'ensemble des contours doivent présenter une continuité identique le long des lignes des images. L'utilisation de cette contrainte impose donc un traitement régulier des lignes des images du couple stéréoscopique à appairier. La contrainte de continuité figurale permet alors de favoriser les appariements qui prolongent les contours le long des lignes des deux images en propageant les hypothèses de mise en correspondance d'une ligne à la suivante : les primitives appariées sur une ligne sont utilisées pour confirmer les appariements de la ligne suivante. Il s'agit alors de privilégier, sur la ligne courante, les appariements qui prolongent un contour mis en évidence par les appariements établis durant le traitement des lignes précédentes.

## 1.8 Mise en correspondance

L'étape la plus délicate d'un processus de stéréovision est la mise en correspondance des primitives extraites de deux images stéréoscopiques. Il n'existe pas de méthode universelle qui puisse s'appliquer à tout couple d'images stéréoscopiques, sans tenir compte de leur nature, de leur contenu et de la façon dont elles sont traitées. D'un point de vue algorithmique, les méthodes développées comportent généralement une forme ou une autre de recherche des meilleurs appariements dans un ensemble fini d'appariements possibles. La qualité d'un algorithme de mise en correspondance se mesure au nombre de primitives appariées, qui doit être le plus élevé possible, et au taux de faux appariements, qui doit être aussi faible que possible.

Les méthodes de mise en correspondance peuvent être classées en deux catégories : les méthodes surfaciques et celles à base de primitives.

### 1.8.1 Méthodes surfaciques

Les méthodes surfaciques cherchent des mises en correspondance de pixels, et se basent sur une mesure de corrélation pour évaluer la qualité des appariements [145]. Ces méthodes fournissent une carte de disparité dense et sont généralement utilisées pour des applications de reconstruction de surfaces d'objets. Les méthodes surfaciques ne sont pas adaptées aux images contenant des textures répétitives ou des zones de niveaux de gris homogènes [132]. En effet, dans le cas d'images contenant des textures répétitives, la fonction de similarité présente plusieurs minima locaux. Dans le cas d'images contenant des zones de niveaux de gris homogènes, la fonction de similarité ne présente pas de pic significatif. Dans les deux cas, il est difficile d'exploiter la fonction de similarité. De plus, à cause de la combinatoire des tests à effectuer, les méthodes surfaciques nécessitent une puissance de calcul importante. Cependant, sachant que les opérations à réaliser sont identiques lors du balayage des images à mettre en correspondance, le traitement peut être décomposé et implanté en parallèle sur une architecture de calcul dédiée [132].

### 1.8.2 Méthodes à base de primitives

Les méthodes à base de primitives sont généralement utilisées pour localiser des objets dans une scène en utilisant des indices visuels tels que les points d'intérêt, les contours, les régions, etc. L'utilisation de telles primitives permet de réduire la combinatoire du problème de mise en correspondance, et ainsi le coût calculatoire de la recherche des appariements. De plus, le caractère discriminant de ces primitives augmente les chances d'aboutir à de bons appariements grâce aux différents attributs qu'on peut leur associer. Les méthodes à base de primitives fournissent des cartes de disparité éparse. Pour des applications de reconstruction de surfaces d'objets, la carte de disparité éparse peut être complétée pour obtenir une carte de disparité dense en utilisant des techniques d'interpolation [123]. Dans ce cas, le type de primitive utilisé détermine la difficulté et la qualité de la reconstruction finale de la scène.

### 1.8.3 Techniques de mise en correspondance

Cette section présente les principales techniques de mise en correspondance stéréoscopique. D'autres études bibliographiques sont présentées dans [18, 145, 68].

#### 1.8.3.1 Techniques corrélatives

Les techniques corrélatives sont employées généralement pour appairer des primitives de type point. La recherche des appariements est basée sur un critère de ressemblance photométrique. On considère une fenêtre autour du point que l'on veut appairer, puis on cherche dans l'autre image le point dont la fenêtre, de même taille que la précédente, est la mieux corrélée. La contrainte épipolaire est d'abord utilisée afin de réduire le nombre de points candidats à l'appariement. Certaines contraintes globales (unicité, ordre, continuité de la disparité) sont ensuite appliquées pour lever certaines ambiguïtés. Plusieurs auteurs [118, 23, 121, 158] ont utilisé les algorithmes de corrélation sur des primitives particulières (points de contour, points de Moravec, points correspondants aux passages par zéro du Laplacien, etc.). D'autres méthodes utilisant le pixel comme primitive de mise en correspondance sont présentées dans [145, 132].

L'ajustement de la taille de la fenêtre de corrélation est crucial. En effet, il faut essayer de trouver le meilleur compromis entre les contraintes de temps de calcul et la qualité des résultats, sachant que la présence d'occultations conduit souvent à de faux appariements. Certains auteurs proposent alors d'utiliser des fenêtres adaptatives [82, 103, 104, 17, 125], des fenêtres multiples [54, 51] ou encore des fenêtres prédites [116]. Une autre approche consiste à fixer la taille et la forme de la fenêtre de corrélation avec une modulation de l'importance accordée à chaque point de la fenêtre. Cette technique a été utilisée dans [181] pour suivre des points d'intérêt dans une séquence d'images stéréoscopiques.

Les méthodes corrélatives sont simples à mettre en ouvre. Cependant, elles présentent deux limitations principales :

- Elles sont sensibles aux distorsions de perspective. Un effet de perspective peut influencer les coefficients de corrélation soit par la différence de luminosité appa-

rente des surfaces suivant l'angle de prise de vue, soit par le fait qu'on corrèle des fenêtres de même taille alors qu'en réalité deux zones qui se correspondent n'ont pas les mêmes tailles. Pour contourner ce problème, Ohta [124] utilise un critère de pseudo-corrélation.

- Elles ne permettent pas de détecter les occultations. Une méthode corrélative fournit toujours, pour chaque point, un correspondant qui présente la meilleure corrélation, même si elle est faible. Or, lors d'une occultation, un point d'une image peut ne pas avoir de correspondant dans l'autre image. L'utilisation d'un seuil pour ne retenir que les coefficients de corrélation suffisamment élevés peut conduire à éliminer en même temps certains bons appariements.

### 1.8.3.2 Techniques hiérarchiques

Les techniques hiérarchiques, dites aussi pyramidales ou multi-résolutions, sont généralement utilisées pour mettre en correspondance des primitives de type point. Elles commencent par représenter les images avec plusieurs niveaux de résolution, du niveau le plus fin, correspondant à la résolution du capteur, au niveau le plus grossier, correspondant à une basse résolution. Pour construire une pyramide, on réduit une image en moyennant par exemple la valeur des niveaux de gris des pixels d'une fenêtre  $N \times N$ . Il s'agit d'établir, à partir d'une paire d'images stéréoscopiques, deux pyramides d'images. Chaque pyramide contient l'image initiale ainsi que ses différents niveaux de résolution.

A l'issue de cette première étape, les coordonnées de tout point de l'image gauche (resp. droite) haute résolution sont connues dans chaque image de la pyramide gauche (resp. droite). L'étape suivante consiste à mettre en correspondance les couples d'images, du niveau basse résolution vers le niveau haute résolution, en utilisant la contrainte épipolaire et un critère de ressemblance. Un couple de points apparié à un niveau de résolution conduit à la recherche des appariements dans deux fenêtres de points  $N \times N$ , situées sur les images gauche et droite du niveau supérieur. Ce processus est itéré jusqu'au niveau de résolution le plus fin qui correspond aux images initiales. L'intérêt de cette technique est qu'elle permet de réduire considérablement la combinatoire du problème de mise en correspondance.

Parmi les techniques hiérarchiques, on peut se référer à [148, 8, 100, 176, 72, 1, 33, 90]. Dans [148], les auteurs ont proposé d'apparier les points correspondants aux passages par zéro du Laplacien en utilisant les images binaires du Laplacien à trois niveaux de résolution (256x256, 128x128, 64x64). Il faut toutefois noter que ces méthodes présentent le risque de propager les erreurs d'appariement commises à un niveau de résolution vers les niveaux de résolution supérieurs. Pour tenter de résoudre ce problème, certains auteurs ont proposé d'utiliser des pyramides à recouvrement [2] ou une stratégie de recherche plus complexe avec la possibilité d'effectuer des retours en arrière à l'intérieur de la pyramide [100].

### 1.8.3.3 Techniques basées sur la programmation dynamique

Le problème de l'appariement peut être vu comme un problème de recherche d'un chemin optimal dans un graphe dont les noeuds sont les couples candidats à l'appariement. Il s'agit au départ d'utiliser la contrainte épipolaire et de considérer à chaque fois deux droites épipolaires conjuguées. On établit ensuite un graphe représentant tous les appariements possibles. Ce graphe peut être matérialisé par un tableau à deux dimensions, où chaque dimension représente les primitives d'une droite épipolaire, classées, de gauche à droite, dans le même ordre que celui de leurs positions dans cette droite. A chaque noeud du graphe correspond un élément du tableau qui représente un coût, déterminé à partir d'une mesure de ressemblance, indiquant la qualité de l'appariement de deux primitives. Un chemin du graphe consiste en une séquence de noeuds, construite en parcourant les droites épipolaires de gauche à droite. Le coût d'un chemin n'est autre que la somme des coûts des sommets qui le constituent. L'objectif est donc de trouver dans le graphe un chemin de coût optimal. La nature de la recherche d'un chemin optimal selon le principe de la programmation dynamique impose deux contraintes :

- L'ordre des primitives homologues doit être respecté le long des deux droites épipolaires (contrainte d'ordre).
- Chaque primitive ne doit intervenir qu'une seule fois (contrainte d'unicité).

Il est possible qu'il y ait dans un chemin une ou plusieurs transitions horizontales ou verticales entre les éléments du tableau représentant le graphe. Cette situation se présente quand une ou plusieurs primitives, gauches ou droites, n'ont pas de correspondant. Les transitions entre deux couples de primitives appariées sont diagonales.

Plusieurs auteurs [6, 11, 124] ont utilisé la programmation dynamique pour appairer des images stéréoscopiques en choisissant les points de contour comme primitives. La recherche des correspondances est faite entre les droites épipolaires conjuguées. Outre les contraintes d'unicité et d'ordre, Ohta [124] complète cette recherche en prenant en compte la contrainte de continuité figurale des contours d'une ligne image à l'autre (cf. section 1.7.2.4). D'autres auteurs [99, 175] ont utilisé la programmation dynamique pour appairer des segments de droite. Dans [12], les auteurs ont utilisé la programmation dynamique pour appairer ce qu'ils ont appelé les "déclivités" [137, 111]. Une "déclivité" est définie par un ensemble de pixels limité par deux extrema locaux successifs de la fonction niveau de gris. D'autres auteurs [14, 15] ont utilisé la programmation dynamique pour effectuer la mise en correspondance dense, c'est à dire en utilisant les pixels comme primitives d'appariement.

L'utilisation de la programmation dynamique fournit généralement de bons résultats, mais au prix d'un grand nombre de calculs qui rend les algorithmes plus ou moins lents suivant la complexité des images et la puissance du matériel disponible. Cette technique présente le risque de propager les erreurs d'appariement car les chemins sont construits pas à pas en examinant les sommets d'une manière locale [18]. Il faut également noter que la contrainte d'ordre, imposée impérativement par la programmation dynamique, limite son utilisation à des images où l'ordre des primitives est respecté dans la plupart des cas, comme les images aériennes par exemple [89].

#### 1.8.3.4 Techniques d'étiquetage par relaxation

Le principe de la relaxation, ou étiquetage probabiliste itératif, consiste à étiqueter l'ensemble des données à traiter, c'est à dire à associer des étiquettes à des objets avec des coefficients de confiance. On part souvent d'un étiquetage ambigu, ce qui signifie que plusieurs étiquettes peuvent correspondre à priori à un même objet. Pour réduire l'am-

biguïté, on est amené à faire coopérer les objets entre eux pour modifier les coefficients de confiance attribués aux associations objet-étiquette à l'aide d'un algorithme itératif de relaxation. On suppose que l'étiquetage d'un objet est d'autant plus satisfaisant et fiable qu'il est cohérent avec l'étiquetage de ses voisins. Dans ce but, on définit un critère global que l'on doit optimiser à chaque itération. Pour plus de détails sur ces techniques de relaxation, on peut se référer à [36, 63].

Le problème de la mise en correspondance peut être vu comme un problème d'étiquetage. Deux formulations peuvent être utilisées :

- Les primitives de l'image de référence sont les objets, les étiquettes sont les primitives de l'autre image.
- Les objets sont les paires de primitives appartenant aux deux images et il n'y a que deux étiquettes : 1 (appariement) ou 0 (pas d'appariement).

Supposons qu'on utilise la première formulation. Initialement, un ensemble d'appariements entre les primitives gauches et droites qui définissent des associations objet-étiquette est sélectionné en respectant certaines contraintes, comme la contrainte épipolaire. Une mesure de confiance est ensuite affectée à chaque appariement. L'étape suivante est destinée à comparer les appariements de primitives voisines en utilisant des relations découlant des contraintes géométriques et/ou physiques. Les appariements ainsi que leurs coefficients de confiance sont ensuite modifiés itérativement pour réduire les inconsistances d'appariement entre les primitives voisines. Le processus itératif s'arrête quand les coefficients de confiance ne varient plus d'une itération à la suivante.

De nombreux auteurs ont proposé la relaxation comme méthode de mise en correspondance. Un algorithme "générique" est présenté dans [7]. L'algorithme de Marr et Poggio [106] est initialisé avec tous les appariements qui satisfont la contrainte épipolaire, auxquels sont associés des poids. A chaque itération et pour chaque appariement, un processus inhibiteur réduit le poids des appariements conflictuels, et un processus excitateur augmente le poids des appariements de disparités voisines. Au cours de la phase itérative, les appariements de poids insuffisants sont éliminés. Grimson [62] a proposé d'apparier les points correspondants aux passages par zéro du Laplacien, en combinant la relaxation et une approche multi-résolution : les appariements obtenus à une échelle grossière sont

utilisés pour résoudre les conflits d'appariement à une échelle plus fine. Dans [135, 134], les auteurs ont proposé un algorithme pour appairer les points de contour. L'algorithme est initialisé en utilisant les contraintes épipolaire et d'orientation. Ensuite, les contraintes d'unicité et de continuité de la disparité sont imposées progressivement pendant la phase itérative. L'originalité de cette approche réside dans la formulation de la contrainte de continuité de la disparité qui fait intervenir la notion de gradient de disparité limite entre deux couples de points de contour appariés. D'autres auteurs [110, 102] ont proposé d'appairer des segments de droite provenant d'une approximation polygonale des contours. Les contraintes épipolaire et d'orientation sont imposées initialement. L'algorithme de relaxation utilise progressivement les contraintes d'unicité et de continuité de la disparité pour réduire les ambiguïtés. Nasrabadi [117] a utilisé la relaxation pour mettre en correspondance des segments incurvés. Pour initialiser son algorithme, il a appliqué la contrainte épipolaire aux centres de courbure des segments. D'autres méthodes sont présentées dans [129, 136, 27].

De manière générale, les méthodes basées sur la relaxation donnent de bons résultats grâce à la prise en compte globale de l'environnement sous forme de contraintes. La mise en correspondance est d'autant plus fiable que les primitives sont de haut niveau [109]. En revanche, la fiabilité des appariements est conditionnée par la manière dont les primitives sont obtenues. L'inconvénient majeur de la relaxation est lié à son caractère exhaustif qui nécessite un grand nombre de calculs pour déterminer les coefficients de confiance et les remettre à jour à chaque itération.

### 1.8.3.5 Techniques par isomorphisme de graphes

Les techniques basées sur l'isomorphisme de graphes font appel à une description structurelle des images. Chaque image est décrite au moyen d'un graphe relationnel où les noeuds correspondent aux primitives extraites de l'image et les arcs, ou arêtes, aux relations entre ces primitives. Ainsi, le problème de mise en correspondance est ramené à un problème d'isomorphisme de graphes. En d'autres termes, on est amené à trouver le meilleur appariement possible entre les noeuds et les arcs de deux graphes. Dans la pratique, les deux graphes ne sont pas identiques à cause des occultations, des effets de

perspective, du non respect de la contrainte d'ordre, etc. [72]. Tenant compte de ces aspects, le problème d'appariement devient un problème d'isomorphisme de sous-graphes maximaux. Il s'agit alors de chercher tous les sous-graphes qui se correspondent et sélectionner ensuite le couple de sous-graphes qui met en correspondance le plus grand nombre de noeuds.

Skordas [150] a utilisé cette technique pour appairer des segments de droite entre lesquels différentes relations comme "à gauche de", "à droite de", "colinéaire avec" ou "a une jonction commune avec" sont définies. Pour chercher le meilleur appariement entre les deux graphes représentant les deux images à appairer, un graphe d'association, ou de correspondance, est construit de la manière suivante. Des appariements potentiels sont établis en respectant les contraintes épipolaire et d'orientation. Les noeuds du graphe d'association, représentant ces appariements, sont reliés entre eux dès qu'il y a une compatibilité entre les appariements correspondants. Cette compatibilité est établie à partir des contraintes d'ordre, d'unicité et de continuité figurale ainsi que d'autres propriétés projectives. L'étape suivante consiste à chercher les ensembles de noeuds mutuellement compatibles, formant ainsi des cliques. La clique maximale correspondant au plus grand nombre d'appariements est retenue comme solution [71].

Shen [148] et Zhao [180] ont proposé d'appairer les graphes pyramidaux des images binaires du Laplacien. Dans [29], les auteurs représentent les images à appairer par des graphes relationnels où les noeuds correspondent aux régions extraites des deux images et les arcs aux relations d'adjacence. Après l'établissement d'un ensemble d'appariements potentiels, en utilisant des critères géométrique et photométrique, l'isomorphisme est optimisé en propageant les appariements initiaux à l'aide de relations de voisinage existantes entre les deux graphes. La même technique a été utilisée dans [10] pour appairer les régions extraites d'images stéréoscopiques. Pour modéliser les images sous forme de graphes, l'auteur a proposé d'utiliser des relations topologiques définies à partir des positions des rectangles exinscrits aux régions. Ces relations sont du type : "au dessus de", "au dessous de", "à gauche de", "à droite de", "masquée par" ou "adjacente à".

### 1.8.3.6 Techniques par prédiction et vérification d'hypothèses

Le processus de mise en correspondance utilisant des concepts de prédiction et vérification d'hypothèses se décompose en trois étapes :

1. **Prédiction d'hypothèses** : Il s'agit d'établir un ensemble d'appariements à partir de contraintes locales, telles que la contrainte épipolaire et la contrainte d'orientation, et de critères de ressemblance géométriques et/ou photométriques.
2. **Propagation d'hypothèses** : Le but de cette deuxième étape est d'étendre le nombre d'appariements en propageant aux primitives voisines chacune des hypothèses émises dans la première étape. La propagation des appariements est effectuée d'une manière récursive en utilisant des contraintes globales telles que l'ordre, l'unicité et la continuité de la disparité.
3. **Vérification d'hypothèses** : Cette phase est destinée à comparer les appariements obtenus par la propagation des hypothèses. Les appariements retenus sont ceux qui correspondent au plus grand nombre d'hypothèses compatibles vis-à-vis d'une ou de plusieurs contraintes globales.

Cette technique a été utilisée par Ayache pour appairer des segments de droite [5]. Une transformation géométrique rigide permettant de passer d'une image à l'autre est déduite d'un certain nombre d'hypothèses générées en appariant, par ressemblance, un ou plusieurs segments, sélectionnés parmi les plus longs d'une image, à des segments de l'autre image. Par propagation, on teste si cette transformation permet aussi d'appairer les segments voisins des premiers. Le nombre de segments ainsi appariés lors de chaque propagation confirme ou infirme la validité de cette transformation. D'autres méthodes semblables sont décrites dans [4, 37]. On peut citer aussi les travaux de Krotkov qui intègre cette technique en appariant des segments dans un système stéréoscopique actif, dont on peut contrôler de nombreux paramètres (mise au point, orientation, zoom, etc.) [91]. Worobel procède par prédiction et vérification d'hypothèses pour appairer les graphes d'adjacence des régions extraites des images stéréoscopiques [172]. La mise en correspondance des régions est guidée par l'appariement de leurs contours en utilisant la programmation dynamique. Pour appairer les régions, Benamrane modélise les images

sous forme de graphes relationnels [10]. Après la génération des hypothèses, leur propagation et leur validation sont réalisées par la définition d'une compatibilité sur des relations topologiques entre les régions.

### 1.8.3.7 Techniques par réseaux de neurones

De nombreux auteurs ont utilisé les réseaux de neurones pour traiter le problème de la mise en correspondance stéréoscopique.

Dès 1976, Marr et Poggio ont développé un algorithme itératif faisant coopérer les contraintes d'unicité et de continuité de la disparité [106]. Ils ont montré que cet algorithme peut être implanté en utilisant un réseau de neurones sélectifs à différentes combinaisons de coordonnées spatiales et de disparités rétiniennes, mais sans aucun apprentissage incrusté. Les neurones du réseau sont connectés de telle sorte qu'il y a excitation entre les neurones correspondant à des appariements qui représentent des disparités voisines et inhibition entre les neurones correspondant à des appariements conflictuels.

Khotanzad et al. ont proposé un réseau de neurones multi-couches pour effectuer la mise en correspondance des pixels des images stéréoscopiques [86, 87, 85]. L'apprentissage supervisé du réseau est assuré par l'algorithme de rétro-propagation du gradient, en utilisant les contraintes d'unicité et de continuité de la disparité. Les cartes de disparité présentées à l'entrée du réseau respectent la contrainte épipolaire. Dans cette méthode, seule la contrainte épipolaire est appliquée explicitement. Les contraintes d'unicité et de continuité de la disparité sont codées sur le réseau de neurones grâce à l'apprentissage.

Pajares et al. ont proposé d'apparier des segments de droite en utilisant un réseau de neurones basé sur une technique d'auto-organisation [34, 128]. Dans leur approche, chaque segment de droite est caractérisé par quatre attributs (l'amplitude du gradient, la direction du gradient, la variance et le Laplacien). Chaque paire de primitives est ainsi caractérisée par un point dans un espace à 4 dimensions dont les coordonnées sont les différences entre les valeurs des attributs associés aux primitives. Les auteurs ont constaté que les bons appariements sont représentés par un nuage de points autour d'un point central. Pour estimer le centre du nuage, le réseau de neurones est soumis à un apprentissage non supervisé par un algorithme d'auto-organisation. Il s'agit ensuite de chercher les paires

de primitives qui minimisent la distance de Mahalanobis entre les points caractérisant les paires de primitives et le centre du nuage. Il faut noter que cette méthode effectue une recherche locale des appariements en utilisant un critère de ressemblance entre les primitives. Elle ne fait appel à aucune contrainte globale pour vérifier les ambiguïtés entre les appariements obtenus.

Pour appairer les pixels des images stéréoscopiques, Wang et Hsiao ont proposé une méthode basée sur l'utilisation de deux réseaux de neurones [169]. Chaque pixel est caractérisé par cinq attributs (l'intensité sur une fenêtre, le module du gradient, l'orientation, l'abscisse, l'ordonnée). Le premier réseau utilise un apprentissage compétitif basé sur le principe de "vitality conservation" [168]. Il sert à regrouper les pixels qui ont les mêmes caractéristiques dans les deux images. Au cours de l'apprentissage, ce réseau reçoit en entrée l'image gauche et construit les groupes de pixels. En mode de fonctionnement, il reçoit en entrée l'image droite et associe chaque pixel à l'un des groupes obtenus au cours de la phase d'apprentissage. Le deuxième réseau est un réseau multi-couches à rétro-propagation du gradient. Il est utilisé pour réaliser une mise en correspondance locale entre les pixels. Ce réseau reçoit en entrée la différence entre les trois premiers attributs de chaque couple de pixels et fournit une sortie qui représente un degré de confiance de mise en correspondance. Les deux réseaux sont ensuite utilisés pour mettre en oeuvre l'algorithme de mise en correspondance selon le schéma suivant. Dans une première étape, le premier réseau est utilisé pour mettre en correspondance les groupes de pixels gauches et droits. Cette étape permet de limiter l'espace de recherche des bons appariements au niveau des pixels. Dans une deuxième étape, on considère deux groupes de pixels qui se correspondent et on fait appel au deuxième réseau pour calculer les degrés de confiance d'appariement des couples de pixels. La dernière étape concerne l'application d'un certain nombre de contraintes (épipolaire, géométrique, ordre, continuité de la disparité) pour réduire les ambiguïtés entre les appariements.

Une autre approche consiste à considérer le problème de mise en correspondance comme un problème d'optimisation. Dans un premier temps, une fonction objective, représentant les contraintes du problème, est construite de telle façon que sa valeur minimale corresponde aux bons appariements. Un réseau de neurones de Hopfield est ensuite

utilisé pour minimiser cette fonction objective. Chaque neurone du réseau représente une possibilité d'appariement entre une primitive de l'image gauche et une primitive de l'image droite. Nasrabadi et Choo ont utilisé cette approche pour appairer les points d'intérêt de Moravec [118]. La fonction objective est formulée à partir des contraintes d'unicité et de continuité de la disparité. D'autres méthodes semblables sont décrites dans [74, 75, 114, 115, 120, 156, 76, 127, 98].

### 1.8.3.8 Techniques par algorithmes génétiques

Basés sur la sélection naturelle et les opérations génétiques du monde du vivant, les algorithmes génétiques ou évolutionnistes sont des techniques stochastiques d'optimisation globale [59]. Vaillant et Gueguen ont utilisé les algorithmes génétiques pour appairer des chaînes de points de contour [161]. La fonction objective est construite à partir des contraintes d'unicité, d'ordre et de continuité de la disparité. Tien propose d'utiliser un algorithme génétique pour appairer les segments de droite dans un espace transformé [155]. Dans un premier temps, les segments de droite extraits des images sont transformés en points dans l'espace de Hough. Le problème de mise en correspondance des segments de droite est ainsi ramené à un problème de mise en correspondance de points dans l'espace de Hough. La recherche d'une solution de mise en correspondance, où les appariements respectent au mieux les contraintes de similarité, d'unicité et d'ordre, consiste ensuite à minimiser une fonction objective représentant ces contraintes grâce à un algorithme génétique. Dans [61], les auteurs proposent d'appairer les segments en utilisant un algorithme génétique avec une fonction objective qui utilise la similarité, évaluée à l'aide de la logique floue. Pour mettre en correspondance les pixels, certains auteurs utilisent la corrélation et génèrent plusieurs cartes de disparité en utilisant plusieurs fenêtres de tailles différentes ou plusieurs images correspondant à différents niveaux de résolution [144, 60]. Les algorithmes génétiques sont ensuite utilisés pour chercher la carte de disparité qui optimise une fonction objective représentant les contraintes de similarité et de continuité. D'autres méthodes utilisant les algorithmes génétiques sont présentées dans [35, 177, 105].

### 1.8.3.9 Techniques utilisant le recuit simulé

Le recuit simulé est un algorithme très utilisé pour résoudre des problèmes d'optimisation [88, 92]. Son principe consiste à considérer le système à optimiser comme un système physique défini par ses états et son énergie. L'évolution de cette énergie vers un optimum est assurée en simulant un processus de recuit, partant d'une température élevée pour terminer à une température voisine de zéro.

De nombreux auteurs ont proposé d'utiliser la technique du recuit simulé pour aborder le problème de l'appariement stéréoscopique. On peut citer les travaux de Vinet et Gagalowicz qui ont développé un algorithme d'appariement de régions basé sur la contrainte de cohérence spatiale des adjacences [165]. Un couple  $(R_g, R_d)$  de régions similaires au regard de critères géométrique et photométrique est validé si les régions adjacentes à  $R_g$  sont mises en correspondance avec des régions adjacentes à  $R_d$  et vice-versa. Dans un premier temps, cette contrainte est exprimée sous la forme d'une fonction objective dont le minimum correspond aux bons appariements. La recherche des meilleurs appariements consiste donc à minimiser cette fonction objective en utilisant l'algorithme du recuit simulé. Dans [107], les auteurs ont proposé de mettre en correspondance les pixels des images stéréoscopiques. Le problème d'appariement est également modélisé sous la forme d'un problème de minimisation d'une fonction objective. Celle-ci évalue plusieurs propriétés du monde physique, du capteur et de la scène : unicité-occultation, ressemblance locale, continuité-régularité locale, point de fixation (intersection des axes optiques). Le processus de minimisation combine à la fois l'algorithme du recuit simulé et la méthode d'optimisation Tabou basée sur l'exploitation de l'historique d'une recherche par voisinage [58]. D'autres méthodes de mise en correspondance utilisant le recuit simulé sont présentées dans [130, 76, 66].

### 1.8.3.10 Autres techniques

Certains auteurs considèrent un couple d'images stéréoscopiques comme deux signaux décalés dans le temps. Les images sont dans un premiers temps transformées dans le domaine fréquentiel en utilisant, par exemple, des filtres de Gabor [16]. Les disparités

sont ensuite estimées explicitement par différence de phase [81, 47, 170] ou implicitement par corrélation de phase [48]. Ces techniques permettent d'obtenir des précisions de type subpixel [53]. En revanche, la réponse des filtres en fréquence reste limitée à une petite bande, ce qui restreint l'intervalle des disparités [53, 132]. Pour faire face à ce problème, certains auteurs proposent de traiter les images à différentes résolutions [53], d'autres proposent d'utiliser les dérivées des images au lieu des images elles-mêmes [126]. Il faut noter que les résultats fournis par ces techniques ne sont précis qu'au niveau des bords des objets [132].

Il existe d'autres méthodes de mise en correspondance qui font appel à des techniques statistiques, notamment celles utilisant la théorie de Bayes. Dans [80], les auteurs considèrent que l'image droite peut être obtenue à partir de l'image gauche via une transformation agissant sur les positions des pixels ainsi que sur leurs intensités. Le modèle proposé inclut également les distorsions provenant du bruit et des occultations. Les paramètres de la transformation sont estimés en utilisant une nouvelle formulation Bayésienne, appelée "Kernel Maximum Likelihood", basée sur l'utilisation des noyaux de densité de probabilité [149]. Cheng and Caelli [26] proposent de représenter le problème de mise en correspondance en s'appuyant sur deux modèles : un modèle de données et un modèle à priori. Le modèle de données est introduit pour décrire l'interaction entre l'espace des disparités autorisées (observation) et la carte de disparité vue comme un champ de Markov aléatoire. Le modèle à priori définit les interactions de voisinage à l'intérieur de la carte de disparité. Ces deux modèles sont déterminés par des fonctions statistiques robustes qui sont contrôlées par un certain nombre de paramètres estimés en utilisant les chaînes de Markov – Monte Carlo [57]. La carte de disparité optimale est ensuite obtenue en utilisant une inférence par propagation de croyance [174]. D'autres techniques semblables sont présentées dans [133, 178, 153, 32].

## 1.9 Conclusion

La stéréovision passive est une technique permettant de retrouver la profondeur d'une scène à partir de deux ou plusieurs images prises sous des angles différents. Cette tech-

---

nique se décompose en quatre étapes : le choix de la géométrie du capteur stéréoscopique ainsi que le calibrage, l'extraction des primitives, la mise en correspondance des primitives, puis la reconstruction 3D.

Le problème clé de la stéréovision est la mise en correspondance des primitives extraites des images stéréoscopiques. En plus de sa nature combinatoire, ce problème ne peut pas être traité sans faire référence à la nature des scènes à analyser, au processus d'acquisition des images ainsi qu'à l'objectif visé. Malgré les nombreuses techniques proposées dans la littérature, ce problème est toujours le centre d'intérêt de nombreux chercheurs dans le domaine de la vision artificielle.

Les chapitres qui suivent présentent les travaux auxquels j'ai participé dans le cadre du développement de techniques de mise en correspondance stéréoscopique pour la perception de l'environnement à l'avant d'un véhicule, en particulier pour la détection d'obstacles.

---

---

## CHAPITRE 2

---

Mise en correspondance par réseau de  
neurones de Hopfield

## 2.1 Introduction

Mes premiers travaux de recherche ont été effectués au sein de l'équipe "Image et Décision" du Centre d'Automatique de Lille qui participait au programme PROMETHEUS pour assurer la détection d'obstacles à l'avant d'un véhicule routier par stéréovision linéaire. Cette problématique a été abordée au sein de l'équipe en trois phases. Au début du programme PROMETHEUS, J.-L. Bruyelle a travaillé dans le cadre de sa thèse sur la conception et la réalisation d'un dispositif de prise de vue à base de caméras linéaires [19]. Il a proposé en particulier une technique de calibrage pour ce dispositif et développé les procédures nécessaires à son étalonnage. L'utilisation des caméras linéaires à la place des caméras matricielles traditionnelles permet un gain de temps très substantiel au niveau de l'analyse des couples d'images stéréoscopiques. Suite à ce travail, J.-C. Burie s'est penché dans sa thèse sur l'analyse d'images issues du stéréoscope linéaire afin de développer les algorithmes de détection d'obstacles [21]. Il a proposé une technique d'extraction de points de contour ainsi qu'une méthode séquentielle pour leur mise en correspondance. La troisième phase a commencé dès mon arrivée au sein de l'équipe pour poursuivre les travaux de recherche commencés par J.-L. Bruyelle et J.-C. Burie avec qui j'ai pu travailler durant la dernière année du programme PROMETHEUS. La méthode de mise en correspondance développée par J.-C. Burie ayant montré ses limites, j'ai proposé une nouvelle approche basée sur l'utilisation des réseaux de neurones de Hopfield [139]. Au delà du programme PROMETHEUS, mon travail consistait à développer cette nouvelle approche au sein de l'équipe qui travaillait sur d'autres projets liés à la sécurité dans les transports terrestres. Ces premiers travaux ont été par la suite le point de départ d'autres travaux de recherche au sein de l'équipe "Image et Décision". Ces travaux auxquels j'ai participé ont fait l'objet de deux thèses réalisées au sein de l'équipe.

Cette nouvelle approche consiste à formuler le problème de mise en correspondance sous la forme d'un problème d'optimisation où une fonction objective doit être minimisée. Celle-ci est construite à partir des contraintes du problème de telle sorte que les bons appariements correspondent à son minimum. Le processus de minimisation est assuré par un réseau de neurones de Hopfield. Dans ce réseau, entièrement connecté, chaque neu-

rone représente une possibilité de mise en correspondance entre une primitive de l'image gauche et une primitive de l'image droite.

Bien qu'elle est présentée dans le cadre de la mise en correspondance des points de contour extraits d'images-ligne issues du stéréoscope linéaire (conçu et réalisé par J.-L. Bruyelle dans le cadre du programme PROMETHEUS), cette approche a pu être adaptée et appliquée également aux images matricielles en considérant d'autres types de primitives [42, 130, 143, 164].

Ces premiers travaux ont donné lieu à trois publications dans des journaux, un chapitre de livre, six communications dans des conférences internationales et trois communications dans des conférences nationales.

## 2.2 Stéréovision linéaire

La stéréovision linéaire fait référence à l'utilisation de caméras linéaires qui fournissent des images-ligne représentant des vues en coupe de la scène observée. En effet, le champ de visée de ce type de caméra est réduit à un plan. Un capteur matriciel standard délivre 582 lignes de 750 pixels, soit 436 500 pixels. En comparaison avec un capteur linéaire de 2048 pixels sur sa ligne unique, on obtient 200 fois moins d'information à traiter, tout en ayant une résolution horizontale bien supérieure. L'utilisation des caméras linéaires à la place des caméras matricielles permet de réduire considérablement le temps de traitement des couples d'images stéréoscopiques [79, 78, 31, 19].

Contrairement aux caméras matricielles, la notion de verticale est absente pour les caméras linéaires. Ainsi, pour qu'un couple de caméras linéaires constitue un système stéréoscopique, il faut que leurs plans de visée soient confondus, c'est à dire que les images-ligne représentent la même coupe de la scène observée.

L'équipe "Image et Décision" qui participait au programme PROMETHEUS disposait d'un stéréoscope linéaire qui a été réalisé par J.-L. Bruyelle [19]. Ce dernier a proposé une procédure de calibrage afin d'assurer la coplanéité des plans de visée des deux caméras linéaires et de mettre celles-ci dans une configuration géométrique prédéterminée.

La figure 2.1 illustre la disposition géométrique de deux caméras linéaires constituant

un capteur stéréoscopique. Les axes optiques sont parallèles et écartés d'une distance  $L$ . Les objectifs associés aux deux caméras ont la même distance focale  $f$ . La partie commune aux champs de visée des deux caméras définit ce qu'on appelle le secteur de stéréovision.

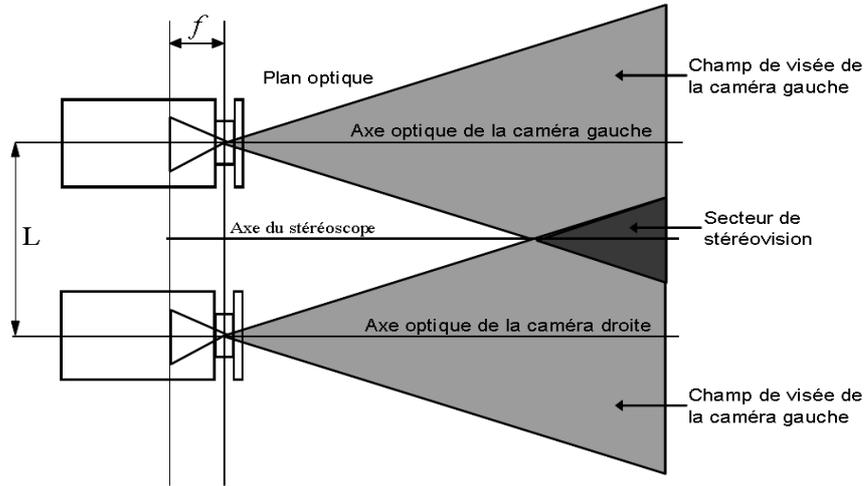


FIG. 2.1 – Géométrie du stéréoscope linéaire.

Avec le modèle du sténopé (cf. Figure 2.2), l'axe  $OX$  représente la ligne joignant les centres optiques des objectifs des deux caméras. L'axe  $OZ$ , perpendiculaire à  $OX$ , est parallèle aux axes optiques des caméras. L'origine  $O$  est située au milieu du segment joignant les centres optiques  $O_g$  et  $O_d$  des objectifs des deux caméras. Soit un point-objet  $P$  repéré dans le système  $\{O, X, Z\}$  par ses coordonnées  $X_P$  et  $Z_P$ . Si ce point est vu par les deux caméras, c'est à dire situé dans le secteur de stéréovision, il produit par projection deux points-image,  $x_g$  sur le capteur linéaire gauche et  $x_d$  sur le capteur linéaire droit. Inversement, les coordonnées du point-objet  $P$  dans le système  $\{O, X, Z\}$  peuvent être retrouvées par triangulation, en exploitant la disparité, ou décalage,  $d = |x_g - x_d|$  entre les points-image  $x_g$  et  $x_d$ . Elles sont exprimées par les équations suivantes :

$$Z_P = \frac{L \cdot f}{d} \quad (2.1)$$

$$X_P = \frac{x_g \cdot Z_P}{f} - \frac{L}{2} = \frac{x_d \cdot Z_P}{f} + \frac{L}{2} \quad (2.2)$$

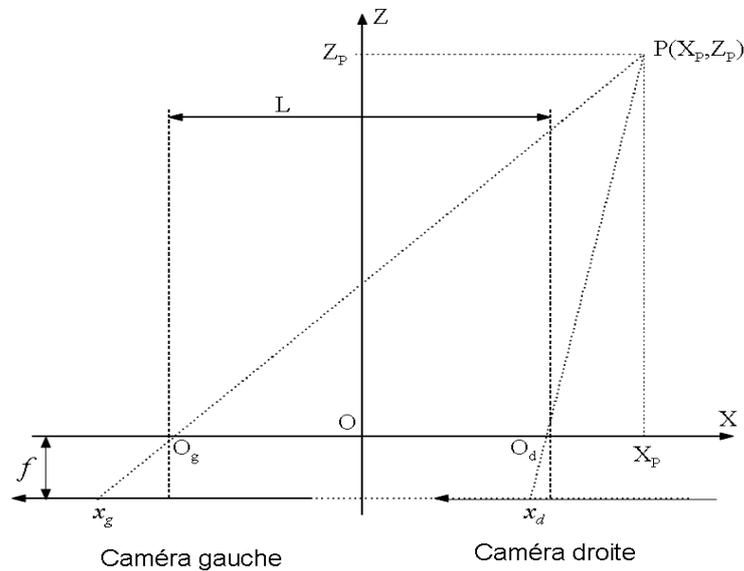


FIG. 2.2 – Modèle du sténopé.

## 2.3 Extraction des primitives

La première étape du processus de stéréovision consiste à extraire de chaque image les primitives à apparier. Comme nous l'avons montré au chapitre 1, la nature de ces primitives et les différents attributs qu'on peut leur associer jouent un grand rôle lors de la phase de mise en correspondance, ainsi que pour la reconstruction de la scène. En vision matricielle, on peut extraire différents types de primitives. Dans le cas de la vision linéaire, ce choix se trouve restreint en raison de la nature mono-dimensionnelle du profil des images-ligne.

La seule possibilité envisageable est la recherche de primitives de type point. Mais cette recherche se trouve limitée puisque les notions de courbure ou de coin n'existent plus dans un espace à une dimension.

L'emploi des images-ligne restreint donc considérablement les points particuliers à

extraire. Ainsi, seuls les points de contour, correspondant aux frontières des différents objets présents dans l'image, peuvent être extraits (cf. Figure 2.3).

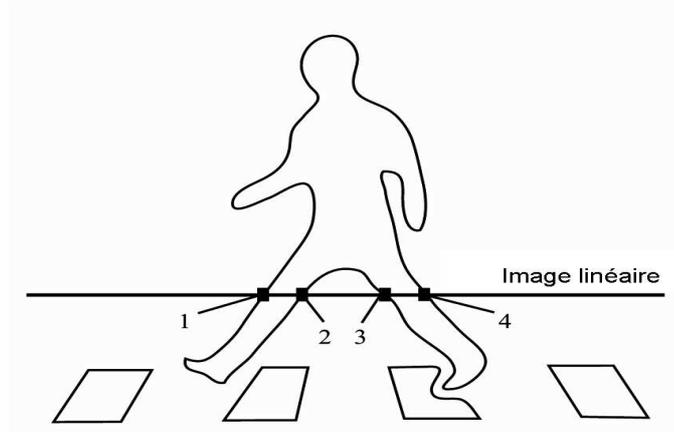


FIG. 2.3 – Primitives de type point de contour en vision linéaire.

La recherche des points de contour consiste à détecter les variations locales de l'intensité de l'image. Il faut toutefois noter que la définition des points de contour en vision linéaire est différente de celle donnée en vision matricielle, puisqu'un point de contour est considéré comme un point unique et non comme une partie d'un contour. Aussi, les propriétés géométriques couramment attribuées aux points de contour en vision matricielle, comme la direction, ne peuvent être définies en vision linéaire.

Les résultats issus de la phase de détection sont souvent inexploitable à cause des fausses détections et des détections manquées. Cependant, il est très important d'extraire de tous les extrema locaux de la fonction dérivée ceux qui correspondent effectivement à des points de contour significatifs de l'image. Les techniques d'extraction les plus utilisées procèdent par un seuillage simple ou un seuillage par hystérésis de l'image gradient [28]. Le seuillage par hystérésis, qui fait appel à la notion de continuité du contour, ne peut pas être utilisé dans le cas des images-ligne où cette notion n'est pas disponible. Le seuillage simple consiste à conserver tous les points de l'image où l'amplitude du gradient est supérieure à un seuil donné. Cependant, il est difficile, voire impossible, de choisir le seuil adéquat permettant à la fois de sélectionner les points de contour significatifs et d'éliminer tous ceux qui ne le sont pas.

Dans sa thèse préparée au sein de l'équipe "Image et Décision" dans le cadre du programme PROMETHEUS, J.-C. Burie a développé une technique originale pour l'extraction des points de contour à partir des images-ligne issues du stéréoscope linéaire. Tout d'abord, la détection des variations locales dans une image-ligne est effectuée grâce aux opérateurs de lissage et de dérivation de Deriche [40]. Pour écarter les inconvénients des méthodes d'extraction par seuillage, il a développé une procédure d'extraction par intervalle [21]. Celle-ci procède en deux étapes. La première consiste à éliminer les extrema générés par le bruit. Dans le cas des images-ligne issues du stéréoscope linéaire, l'expérience a montré que les extrema locaux dus au bruit sont toujours majoritaires, mais leur amplitude est toujours très faible. Pour les éliminer, Burie propose d'appliquer un seuil bas  $S_b$  qui, à l'issue de plusieurs tests, est ajusté à 10% de l'amplitude maximale de l'image gradient. Dans la deuxième étape, il s'agit de sélectionner les extrema significatifs. Dans la pratique, un contour peut donner naissance à plusieurs extrema locaux. En effet, les contours peuvent être plus ou moins larges, et la variation locale du niveau de gris peut être plus ou moins régulière de telle sorte que la dérivée peut présenter plusieurs extrema correspondant néanmoins à un seul contour. En se basant sur cette constatation, la procédure de sélection consiste à découper le signal gradient en une suite d'intervalles adjacents où il conserve un signe constant (cf. Figure 2.4). Dans chaque intervalle, le point ayant la plus forte amplitude est retenu, si et seulement si la valeur absolue de cette amplitude est supérieure au seuil  $S_b$ .

A l'issue de la phase d'extraction, chaque point de contour est caractérisé par trois attributs :

- Sa position dans l'image.
- Le signe de son gradient qui permet d'indiquer s'il s'agit d'un front montant ou d'un front descendant.
- L'amplitude de son gradient.

Dans la poursuite des travaux de l'équipe "Image et Décision", cette technique d'extraction n'a pas été remise en cause. En effet, il nous assure, comme J.-C. Burie l'avait montré, de conserver le maximum de points de contour significatifs tout en éliminant les points

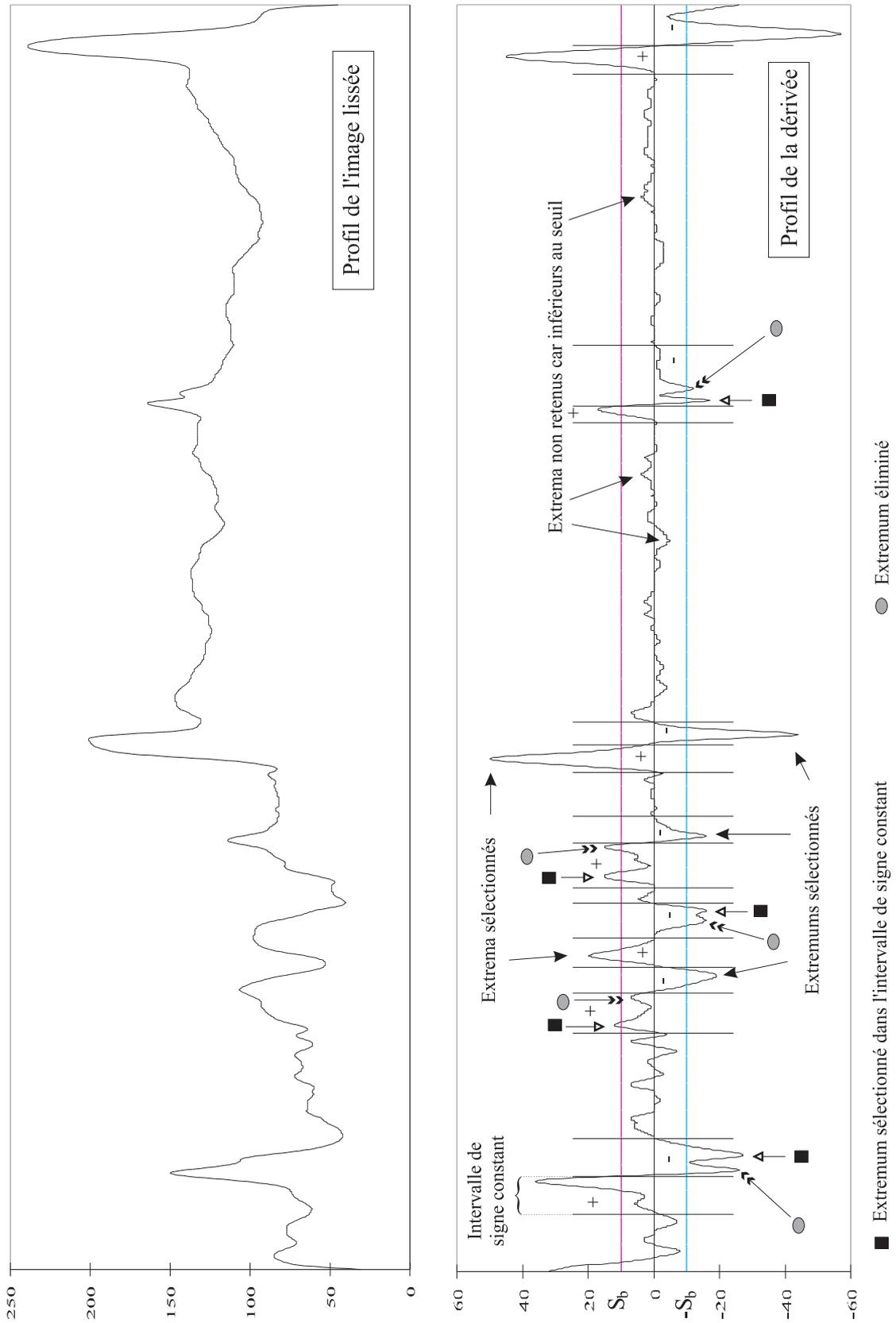


FIG. 2.4 – Extraction des points de contour.

sans intérêt.

## 2.4 Mise en correspondance

Les algorithmes de détection d'obstacles développés au sein de l'équipe "Image et Décision" par J.-C. Burie sont basés sur une procédure de mise en correspondance séquentielle des points de contour extraits d'images-ligne. Utilisant un critère de similarité photométrique, celle-ci procède par une recherche séquentielle des appariements en parcourant les images-ligne du début vers la fin (recherche progressive) et inversement (recherche rétrograde). La liste définitive des couples appariés est l'intersection des listes des couples issues des recherches progressive et rétrograde.

Cette méthode présente deux inconvénients majeurs. Le premier est lié à l'initialisation de l'algorithme de recherche des appariements. En effet, J.-C. Burie part de l'hypothèse que les premiers et les derniers pixels d'un couple d'images-ligne constituent deux couples de points appariés. Le deuxième inconvénient réside dans la nature séquentielle de l'analyse des images-ligne. En effet, la validation d'un couple de points de contour ne dépend que du dernier couple validé. Ainsi, à part l'ordre des appariements qui est intrinsèque au parcours séquentiel des deux images-ligne, aucun critère de compatibilité globale n'est utilisé pour lever les ambiguïtés d'appariement. Par conséquent, si une erreur de mise en correspondance se produit, elle se propage tout au long du processus d'appariement. Afin de diminuer les risques de propagation des erreurs d'appariement, J.-C. Burie a proposé d'utiliser sa méthode séquentielle de manière hiérarchique pour appairer les points de contour à différents niveaux de recherche. Nous allons revenir, dans le chapitre 4, sur cette notion de niveau de recherche que nous avons exploitée dans le cadre d'un travail en commun pour améliorer la cadence des traitements de nos procédures de mise en correspondance. La combinaison de la méthode séquentielle et la stratégie de recherche hiérarchique permet d'obtenir des résultats satisfaisants lorsque les scènes sont simples à analyser, mais se trouve pratiquement inexploitable quand il s'agit d'analyser des scènes complexes. Ici, les mots "simple" et "complexe" expriment le contenu des images-ligne en termes de nombre et de nature des objets présents.

La méthode proposée par J.-C. Burie ayant montré ses limites, j'ai proposé une nouvelle approche de mise en correspondance basée sur l'utilisation des réseaux de neurones de Hopfield [139]. En se plaçant dans une démarche de recherche globale des appariements, le problème de mise en correspondance est formulé sous la forme d'un problème d'optimisation où une fonction objective doit être minimisée. Celle-ci est construite à partir des contraintes du problème de telle sorte que son minimum corresponde à une solution où les appariements respectent au mieux ces contraintes. Le processus de minimisation est assuré par un réseau de neurones de Hopfield. Dans ce réseau, entièrement connecté, chaque neurone représente une possibilité de mise en correspondance entre un point de contour de l'image gauche et un point de contour de l'image droite.

### 2.4.1 Codage du problème

Soit  $G$  (respectivement  $D$ ) la liste des points de contour extraits de l'image gauche (respectivement droite). Le problème de mise en correspondance peut être codé sous la forme de  $N_G \times N_D$  variables binaires d'états, où  $N_G$  (respectivement  $N_D$ ) est le nombre de points de contour de l'image gauche (respectivement droite). Une variable d'état  $E_{gd}$  représente une possibilité de mise en correspondance entre un point de contour de la liste  $G$  et un point de contour de la liste  $D$  :  $E_{gd} = 1$  si les points de contour  $g$  et  $d$  sont appariés,  $E_{gd} = 0$  dans le cas contraire. Ces variables d'états sont groupées sous la forme d'une matrice d'état  $\mathbf{E}$  de taille  $N_G \times N_D$ , appelée matrice de mise en correspondance (cf. Figure 2.5).

5	E <sub>51</sub>	E <sub>52</sub>	E <sub>53</sub>	E <sub>54</sub>	E <sub>55</sub>	E <sub>56</sub>
4	E <sub>41</sub>	E <sub>42</sub>	E <sub>43</sub>	E <sub>44</sub>	E <sub>45</sub>	E <sub>46</sub>
3	E <sub>31</sub>	E <sub>32</sub>	E <sub>33</sub>	E <sub>34</sub>	E <sub>35</sub>	E <sub>36</sub>
2	E <sub>21</sub>	E <sub>22</sub>	E <sub>23</sub>	E <sub>24</sub>	E <sub>25</sub>	E <sub>26</sub>
1	E <sub>11</sub>	E <sub>12</sub>	E <sub>13</sub>	E <sub>14</sub>	E <sub>15</sub>	E <sub>16</sub>

G

D	1	2	3	4	5	6
---	---	---	---	---	---	---

FIG. 2.5 – Matrice de mise en correspondance.

Le problème de mise en correspondance peut être considéré comme un problème de satisfaction de contraintes. Une solution au problème est celle dont l'ensemble des appariements respectent au mieux ces contraintes. Pour quantifier cette optimalité, il suffit de construire une fonction objective  $H$  qui est une application de  $\{0, 1\}^{N_G \times N_D}$  dans  $\mathbb{R}$ . Celle-ci associe à chaque solution un nombre réel qui représente la qualité de la solution ou le degré de compatibilité des appariements vis-à-vis des contraintes imposées. Il s'agit alors de chercher la solution qui correspond au minimum ou au maximum de la fonction objective.

Le problème de mise en correspondance peut être ainsi formulé sous la forme d'un problème d'optimisation où il s'agit de chercher le minimum ou le maximum d'une fonction objective définie à partir des contraintes du problème.

## 2.4.2 Contraintes de mise en correspondance

Les contraintes de mise en correspondance peuvent être locales ou globales.

### 2.4.2.1 Contraintes locales

Les contraintes locales sont exprimées en fonction des attributs des points de contour gauches et droits. Elles sont dites locales car elles sont appliquées à chaque couple de points de contour sans faire intervenir les autres couples. Leur utilisation permet de réduire

la combinatoire du problème de mise en correspondance. Notre procédure de mise en correspondance utilise deux contraintes locales :

- **La contrainte de position** : En stéréovision matricielle, la contrainte épipolaire impose que deux points de contour homologues se trouvent nécessairement sur des droites épipolaires conjuguées. Cette contrainte, qui ne nécessite aucune hypothèse sur la scène observée, dépend uniquement de la géométrie du capteur stéréoscopique. Dans le cas de la stéréovision linéaire, cette contrainte est intrinsèque au capteur stéréoscopique lui-même. En effet, deux images-ligne doivent correspondre à la même coupe de la scène. L'équation 2.2 permet d'écrire (cf. section 2.2) :

$$\frac{x_g \cdot Z_P}{f} - \frac{L}{2} = \frac{x_d \cdot Z_P}{f} + \frac{L}{2} \quad (2.3)$$

On en déduit alors que :

$$x_g - x_d = \frac{L \cdot f}{Z_P} \quad (2.4)$$

Les paramètres  $L$ ,  $f$  et  $Z_P$  étant positifs, on obtient la contrainte  $x_g > x_d$ . Cette contrainte stipule alors que pour qu'un couple de points de contour soit candidat à l'appariement, il faut que la position du point de contour gauche dans l'image gauche soit supérieure à celle du point de contour droit dans l'image droite.

- **La contrainte de signe du gradient** : Il s'agit d'une contrainte photométrique. Elle utilise l'attribut de signe du gradient associé aux points de contour (cf. section 2.3). Elle postule que deux points de contour gauche et droit, engendrés par le même point-objet de la scène, sont caractérisés par des dérivées de la fonction niveau de gris de même signe. Ainsi, pour qu'un couple de points de contour soit candidat à l'appariement, il faut que les points de contour aient le même signe du gradient. Il faut toutefois noter que cette affirmation peut être mise en défaut (cf. Figure 2.6). Le point  $P$  placé à la surface de l'objet est vu par les deux caméras, mais avec des fonds différents. Il est vu par la caméra gauche sur un fond sombre alors qu'il est vu par la caméra droite sur un fond clair. De ce fait, les projections

gauche et droite de ce point génèrent des gradients de signes différents. Cependant, ce type de situation est extrêmement rare dans un contexte routier. En effet, le fond est constitué dans la plupart des cas par le revêtement de la chaussée qui peut être considéré comme homogène puisque la taille des marquages au sol est en général petite par rapport à celle des objets présents sur la route.

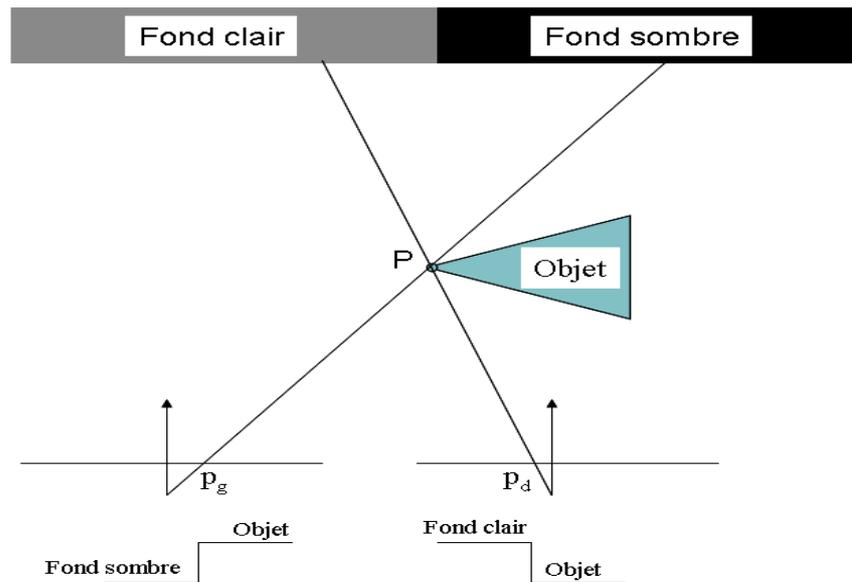


FIG. 2.6 – Mise en défaut de la contrainte de signe du gradient.

#### 2.4.2.2 Contraintes globales

Les contraintes globales sont appliquées en prenant en compte tous les couples de points de contours possibles. Elles permettent de vérifier la compatibilité entre les couples afin de réduire les ambiguïtés de mise en correspondance. Nous utilisons trois contraintes globales :

- **La contrainte d'unicité** : En stéréovision, on peut se trouver dans des situations où un élément est vu par une caméra et pas par l'autre. De telles situations se produisent soit parce que cet élément est occulté par un autre du côté de l'une des deux caméras, soit parce qu'il se trouve hors du champ de vision commun des deux caméras, soit enfin à cause de perturbations telles que les ombres et les reflets. Par conséquent, un point de contour d'une image peut ne pas avoir de correspondant

dans l'autre image. Il en résulte alors une contrainte d'unicité partielle qui stipule qu'un point de contour de l'image gauche a au plus un correspondant dans l'image droite, et inversement. Pour pouvoir l'utiliser, on suppose toujours l'unicité totale, c'est à dire qu'un point de contour d'une image a un correspondant et un seul dans l'autre image.

- **La contrainte d'ordre** : Cette contrainte exprime que les points de contour doivent être appariés en respectant l'ordre de leurs positions dans les deux images. En d'autres termes, deux couples de points de contour  $(g, d)$  et  $(g', d')$  sont compatibles au sens de la contrainte d'ordre, si la position dans l'image gauche de  $g'$  par rapport à  $g$  est la même que celle de  $d'$  par rapport à  $d$  dans l'image droite. Nous avons montré au chapitre 2 que face à certaines situations cette contrainte n'est plus valable. Nous supposons que ces situations sont extrêmement rares compte tenu de la taille des objets rencontrés dans un environnement routier et de la distance à laquelle ils sont détectés.
- **La contrainte de continuité de la disparité** : Cette contrainte est une conséquence directe de l'hypothèse de continuité des surfaces par morceaux : un même objet ne peut présenter de variations brutales de profondeur le long de petites surfaces. La profondeur étant liée à la disparité (cf. Equation 2.1), il en résulte que les disparités doivent varier continûment le long des surfaces des objets, c'est à dire que les points voisins dans l'espace ont des disparités voisines sur les images. Nous utilisons cette contrainte car nous supposons que les objets présents dans un contexte routier ont des formes suffisamment régulières pour qu'il n'y ait pas à leurs surfaces de forte variations de profondeur.

### 2.4.3 Fonction objective

La fonction objective peut être définie selon deux schémas. Le premier consiste à utiliser la fonction objective pour coder toutes les contraintes [140]. L'inconvénient de cette stratégie est qu'elle met en jeu inutilement les couples de points de contour qui ne respectent pas les contraintes locales, c'est à dire les contraintes de position et de signe du gradient. Dans le deuxième schéma, il s'agit de ne considérer que les couples respectant les

contraintes locales et de construire la fonction objective à partir des contraintes globales, c'est à dire les contraintes d'unicité, d'ordre et de continuité de la disparité [141]. C'est ce dernier schéma que nous allons présenter dans ce qui suit.

La fonction objective est définie en ne prenant en compte que les variables d'états qui correspondent aux couples candidats à l'appariement, c'est à dire ceux qui respectent les contraintes locales. Traduisant les contraintes globales, la fonction objective est construite de telle sorte que son minimum corresponde aux bons appariements.

### 2.4.3.1 Fonction objective représentant la contrainte d'unicité

La contrainte d'unicité stipule qu'un point de contour d'une image a un et un seul correspondant dans l'autre image (cf. Section 2.4.2).

Rappelons que la matrice d'état  $\mathbf{E}$  représente tous les appariements possibles entre les points de contours de l'image gauche et ceux de l'image droite (cf. Section 2.4.1). Chaque variable d'état  $E_{gd}$  représente une possibilité de mise en correspondance entre le point de contour  $g$  de l'image gauche et le point de contour  $d$  de l'image droite : si  $E_{gd} = 1$ , le couple candidat  $(g, d)$  est apparié, sinon il ne l'est pas. Une ligne (respectivement colonne) de la matrice d'état représente l'ensemble des appariements potentiels d'un point de contour de l'image gauche (respectivement droite).

Pour satisfaire la contrainte d'unicité, il faut que chaque ligne de la matrice d'état mette en évidence un seul appariement. En d'autres termes, il faut qu'il y ait dans chaque ligne une seule variable d'état à 1 et les autres à 0. Le même raisonnement s'applique aux colonnes de la matrice d'état.

La fonction objective  $H_u$  représentant la contrainte d'unicité est construite de telle sorte que les solutions ne respectant pas cette contrainte soient défavorisées. Elle s'exprime sous la forme suivante :

$$H_u = \sum_{g \in G} \left( 1 - \sum_{d \in D / (g,d) \in \Omega} E_{gd} \right)^2 + \sum_{d \in D} \left( 1 - \sum_{g \in G / (g,d) \in \Omega} E_{gd} \right)^2 \quad (2.5)$$

où  $\Omega$  représente l'ensemble des couples candidats à l'appariement, c'est à dire les couples de points de contour gauche et droit qui respectent les contraintes locales :

$$\Omega = \{(g, d) \in G \times D / (g, d) \text{ vérifie les contraintes de position et de signe du gradient}\}$$

Le premier (respectivement second) terme de la fonction  $H_u$  sert à défavoriser les situations de non respect de la contrainte d'unicité pour les points de contour de l'image gauche (respectivement droite). La fonction objective  $H_u$  atteint son minimum pour les solutions respectant la contrainte d'unicité.

### 2.4.3.2 Fonction objective représentant la contrainte d'ordre

Soient deux couples de points de contour  $(g, d)$  et  $(g', d')$  respectant les contraintes locales, où  $g$  et  $g'$  (respectivement  $d$  et  $d'$ ) sont des points de contour de l'image gauche (respectivement droite). Ces deux couples sont compatibles vis-à-vis de la contrainte d'ordre si l'ordre des positions des points de contour  $g$  et  $g'$  dans l'image gauche est le même que celui des points de contour  $d$  et  $d'$  dans l'image droite. En d'autres termes, la contrainte est respectée si la position dans l'image gauche de  $g'$  par rapport à  $g$  est la même que celle de  $d'$  par rapport à  $d$  dans l'image droite.

La fonction objective  $H_o$  représentant la contrainte d'ordre est construite de telle façon que les solutions mettant en défaut cette contrainte soient défavorisées. Elle peut s'écrire sous la forme suivante :

$$H_o = \sum_{(g,d) \in \Omega} \sum_{\substack{(g',d') \in \Omega \\ g' \neq g \text{ et } d' \neq d}} O_{gdg'd'} E_{gd} E_{g'd'}. \quad (2.6)$$

Le coefficient  $O_{gdg'd'}$  représente une pénalité qui intervient au cas où la contrainte n'est pas respectée par les couples  $(g, d)$  et  $(g', d')$ . Il s'exprime de la manière suivante :

$$O_{gdg'd'} = |\text{signe}(x_g - x_{g'}) - \text{signe}(x_d - x_{d'})|$$

$x_g$ ,  $x_{g'}$ ,  $x_d$  et  $x_{d'}$  sont respectivement les positions des points de contour  $g$ ,  $g'$ ,  $d$  et  $d'$  dans les images gauche et droite. La fonction signe est définie comme suit :

$$\text{signe}(a) = \begin{cases} 1 & \text{si } a > 0 \\ 0 & \text{sinon} \end{cases}$$

La fonction objective  $H_o$  atteint son minimum pour les solutions respectant la contrainte d'ordre.

### 2.4.3.3 Fonction objective représentant la contrainte de continuité de la disparité

La contrainte de continuité de la disparité stipule que les appariements de points de contour voisins doivent donner des disparités voisines. La fonction objective  $H_c$  correspondante est construite pour favoriser les appariements ayant des disparités voisines. Elle peut s'exprimer sous la forme suivante :

$$H_c = - \sum_{(g,d) \in \Omega} \sum_{\substack{(g',d') \in \Omega \\ g' \neq g \text{ et } d' \neq d}} S_{gdg'd'} E_{gd} E_{g'd'}. \quad (2.7)$$

Le coefficient  $S_{gdg'd'}$  représente une mesure de compatibilité entre les couples candidats  $(g, d)$  et  $(g', d')$  vis-à-vis de la contrainte de continuité de la disparité. Il peut s'écrire de la façon suivante (cf. Figure 2.7) :

$$S_{gdg'd'} = S(X_{gdg'd'}) = \frac{2}{1 + e^{\alpha(X_{gdg'd'} - \theta)}} - 1 \quad (2.8)$$

où  $X_{gdg'd'}$  est la valeur absolue de la différence des disparités des couples  $(g, d)$  et  $(g', d')$  :

$$X_{gdg'd'} = |(x_g - x_d) - (x_{g'} - x_{d'})|$$

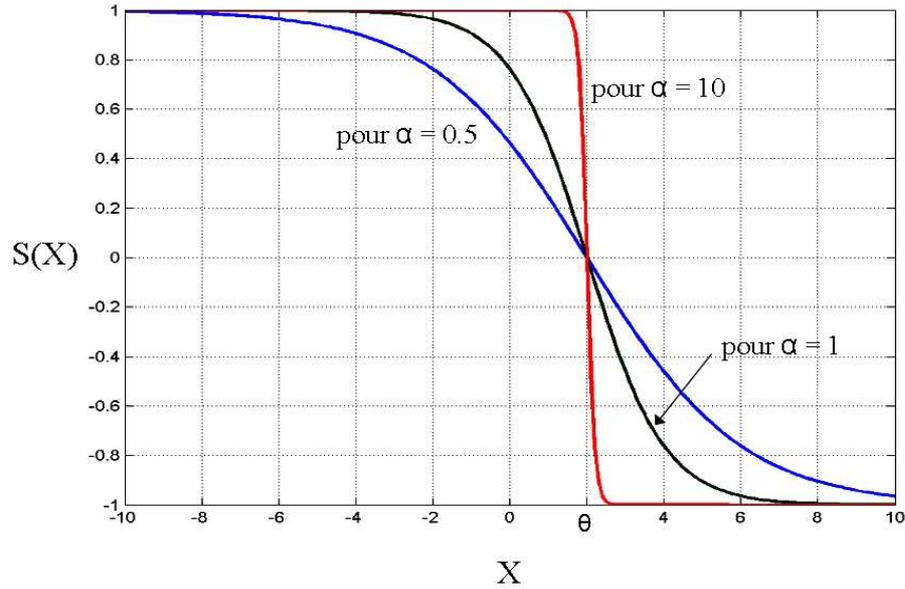


FIG. 2.7 – Fonction de mesure de compatibilité pour différentes valeur de  $\alpha$ .

La mesure de compatibilité  $S(X_{gdg'd'})$  varie continûment entre  $-1$  et  $1$ . Elle est d'autant plus forte, c'est à dire proche de  $1$ , que les disparités des couples  $(g, d)$  et  $(g', d')$  sont voisines. Inversement, elle est d'autant plus faible, c'est à dire proche de  $-1$ , que les disparités des couples  $(g, d)$  et  $(g', d')$  sont éloignées.

Le paramètre  $\theta$ , qui contrôle la position où la fonction  $S(X_{gdg'd'})$  s'annule, permet une certaine tolérance afin de prendre en compte les effets du bruit et les distorsions. Il est ajusté de telle sorte qu'une compatibilité maximale, c'est à dire proche de  $1$ , soit obtenue pour les paires de couples qui correspondent à des valeur de  $X_{gdg'd'}$  voisines de  $0$ , et qu'une compatibilité minimale, c'est à dire proche de  $-1$ , soit obtenue pour les paires de couples qui correspondent à des valeurs élevées de  $X_{gdg'd'}$ . Le paramètre  $\alpha$  contrôle la pente de la fonction  $S(X_{gdg'd'})$  en  $\theta$ . Plus  $\alpha$  est grand, plus la pente est forte, c'est à dire plus on tend vers une fonction binaire de valeurs  $-1$  et  $1$ . Inversement, plus  $\alpha$  est proche de  $0$ , plus la pente est faible, c'est à dire plus le passage entre  $-1$  et  $1$  est lisse.

La fonction objective  $H_c$  atteint son minimum pour les solutions respectant au mieux la contrainte de continuité de la disparité.

### 2.4.3.4 Fonction objective totale

En combinant les contraintes globales, la fonction objective finale pour notre problème de mise en correspondance prend la forme suivante :

$$\begin{aligned}
H_{objective} = & K_u \left[ \sum_{g \in G} \left( 1 - \sum_{d \in D / (g,d) \in \Omega} E_{gd} \right)^2 + \sum_{d \in D} \left( 1 - \sum_{g \in G / (g,d) \in \Omega} E_{gd} \right)^2 \right] \\
& + K_o \sum_{(g,d) \in \Omega} \sum_{\substack{(g',d') \in \Omega \\ g' \neq g \text{ et } d' \neq d}} O_{gdg'd'} E_{gd} E_{g'd'} \\
& - K_c \sum_{(g,d) \in \Omega} \sum_{\substack{(g',d') \in \Omega \\ g' \neq g \text{ et } d' \neq d}} S_{gdg'd'} E_{gd} E_{g'd'}
\end{aligned} \tag{2.9}$$

où  $K_u$ ,  $K_o$  et  $K_c$  représentent respectivement les coefficients de pondération des termes correspondant aux contraintes d'unicité, d'ordre et de continuité de la disparité.

Le problème est à présent formulé sous la forme d'un problème de minimisation d'une fonction objective représentant toutes les contraintes imposées. Dans ce qui suit, nous présentons la procédure neuronale que nous avons développée pour minimiser cette fonction objective.

### 2.4.4 Procédure neuronale pour la minimisation de la fonction objective

Notre procédure de minimisation de la fonction objective  $H$  est basée sur l'utilisation d'un réseau de neurones de Hopfield [69]. Il s'agit d'un réseau où chaque neurone est connecté à tous les autres, même avec lui-même (cf. Figure 2.8). En plus des entrées issues des autres neurones, chaque neurone possède une entrée provenant de l'extérieur. Chaque neurone est caractérisé par un état qui est transmis à tous les autres neurones via des connexions pondérées.

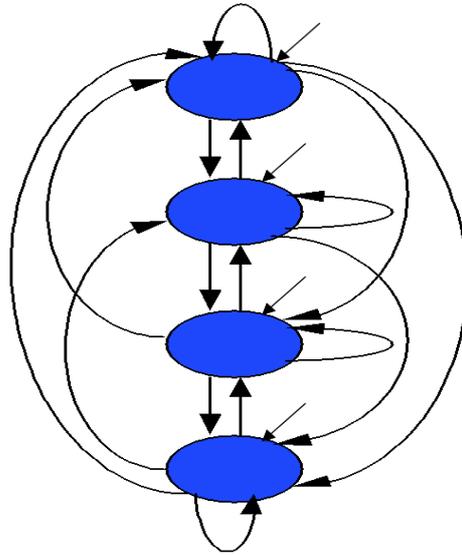


FIG. 2.8 – Réseau de neurones entièrement connecté.

Soit un réseau de Hopfield à  $N$  neurones (cf. Figure 2.9). Soit  $W$  sa matrice des poids des connexions,  $\vec{I} = (I_1, I_2, \dots, I_N)^t$  son vecteur d'entrées externes et  $\vec{V} = (V_1, V_2, \dots, V_N)^t$  son vecteur d'état, composé des états élémentaires des neurones. Le réseau peut être caractérisé par une fonction d'énergie, définie par [69] :

$$\begin{aligned}
 H_{net} &= -\frac{1}{2} \cdot \vec{V}^t \cdot \mathbf{W} \cdot \vec{V} - \vec{I} \cdot \vec{V} \\
 &= -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N W_{ij} \cdot V_i \cdot V_j - \sum_{i=1}^N I_i \cdot V_i
 \end{aligned} \tag{2.10}$$

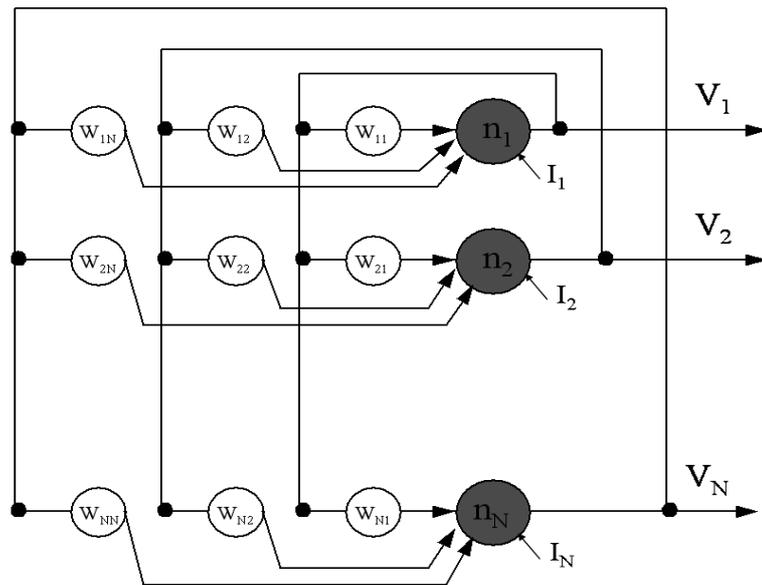


FIG. 2.9 – Réseau de neurones de Hopfield avec sa matrice des poids de connexions et son vecteur d'entrées externes.

Selon le mode de fonctionnement utilisé, c'est à dire la manière selon laquelle les neurones changent leurs états, on distingue trois modèles de réseaux de Hopfield [139] :

- **Le modèle binaire déterministe** : Dans ce modèle, l'état des neurones est binaire. Chaque neurone calcule son potentiel d'entrée, qui est la somme pondérée des sorties provenant des autres neurones augmentée de son entrée externe. Le neurone change ensuite son état en comparant son potentiel d'entrée à un seuil donné.
- **Le modèle binaire stochastique** : Dans ce modèle, l'état des neurones est binaire. Sa mise à jour est basée sur une règle stochastique utilisant le principe du recuit simulé.
- **Le modèle continu** : Dans ce modèle, l'état des neurones varie continûment dans l'intervalle  $[0, 1]$ . La mise à jour de l'état des neurones est régie par une équation différentielle où celui-ci est considéré comme une fonction continue de la variable temporelle, elle aussi continue.

Sous certaines conditions, chacun de ces modes de fonctionnement induit un processus de minimisation de la fonction d'énergie du réseau [139]. En partant d'un état initial,

le réseau évolue jusqu'à ce qu'il atteigne un état stable, c'est à dire un état où les neurones ne changent plus d'état. Cet état correspond alors à un minimum de la fonction d'énergie.

Les modèles binaire et continu induisent un processus de minimisation locale. Le modèle stochastique permet une recherche globale du minimum, mais avec une convergence plus lente due à l'utilisation du recuit simulé. Par rapport au modèle binaire, le modèle continu présente les avantages suivants [73] :

- Le modèle continu tend à éviter les oscillations entre les états stables durant son évolution. De ce fait, la convergence de ce modèle est plus rapide que celle du modèle binaire.
- Les solutions fournies par le modèle continu sont meilleures que celles produites par le modèle binaire car les vallées du paysage énergétique sont généralement plus larges.

Dans le cadre de notre application, le modèle stochastique ne peut être utilisé car il implique une très grande quantité de calcul. Nous avons donc retenu le modèle continu.

Pour pouvoir minimiser notre fonction objective par un réseau de neurones de Hopfield, il suffit de mettre celle-ci sous la forme de la fonction d'énergie du réseau. On détermine ensuite, par identification, la matrice des poids de connexions ainsi que le vecteur d'entrées externes. Cependant, pour que le processus de minimisation converge, il faut disposer d'une matrice de poids symétrique [139].

#### 2.4.4.1 Construction du réseau

En se basant sur la formulation exposée dans la section 2.4.1, la représentation du problème de mise en correspondance par un réseau de neurones de Hopfield est immédiate. Soient deux images gauche et droite ayant respectivement  $N_G$  et  $N_D$  points de contour. Par analogie avec la matrice de mise en correspondance introduite dans la section 2.4.1, il suffit de considérer un réseau de  $N_G \times N_D$  neurones où chacun d'eux représente une possibilité de mise en correspondance entre un point de contour de l'image gauche et un point de contour de l'image droite (cf. Figure 2.10).

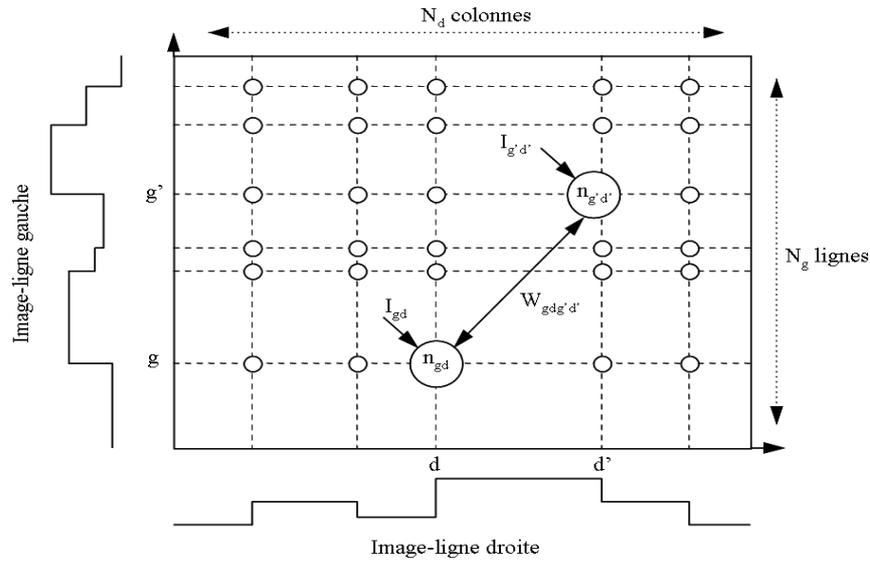


FIG. 2.10 – Réseau de neurones de Hopfield représentant tous les couples possibles.

Plus explicitement, un neurone  $n_{gd}$  représente une possibilité d'appariement entre le point de contour  $g$  de l'image gauche et le point de contour  $d$  de l'image droite. L'état  $E_{gd}$  du neurone  $n_{gd}$  indique si le couple  $(g, d)$  est apparié ou non (cf. Section 2.4.1) : si  $E_{gd} = 1$  alors le couple  $(g, d)$  est apparié, sinon il ne l'est pas.

Tel qu'il est défini, le réseau représente tous les couples de points de contour possibles. Or, nous sommes partis du fait qu'il ne faut considérer que les couples candidats à l'appariement, c'est à dire ceux qui respectent les contraintes locales. Pour simplifier les notations, nous gardons la même représentation, mais sans tenir compte des neurones qui correspondent aux couples ne respectant pas les contraintes locales (cf. Figure 2.11). Bien entendu, ces neurones disparaissent au niveau de l'implantation.

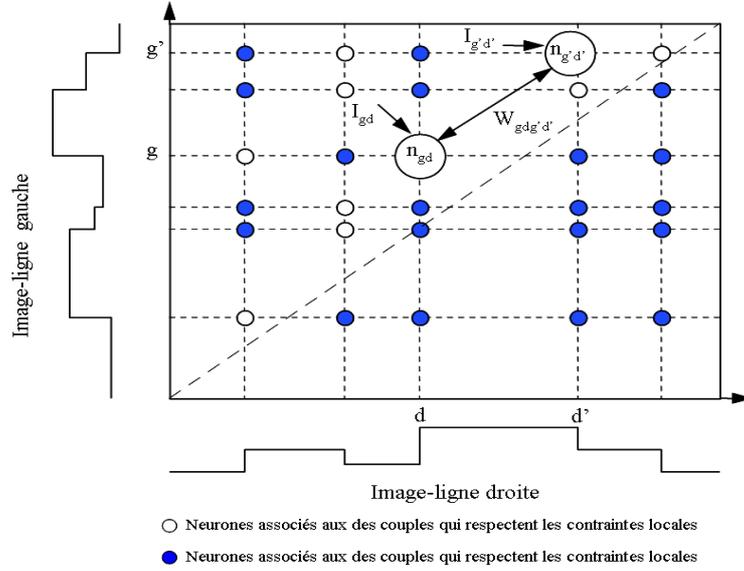


FIG. 2.11 – Réseau de neurones de Hopfield ne représentant que les couples candidats à l'appariement.

Pour compléter la construction du réseau, il faut maintenant déterminer la matrice des poids des connexions  $\mathbf{W}$  et le vecteur d'entrées externes  $\vec{I}$ . Pour cela, il suffit de mettre la fonction objective sous la forme de la fonction d'énergie du réseau :

$$H_{net} = -\frac{1}{2} \sum_{(g,d) \in \Omega} \sum_{(g',d') \in \Omega} W_{gdg'd'} \cdot E_{gd} \cdot E_{g'd'} - \sum_{(g,d) \in \Omega} I_{gd} \cdot E_{gd} \quad (2.11)$$

En identifiant l'expression de cette fonction d'énergie  $H_{net}$  (cf. Equation 2.11) et celle de la fonction objective du problème  $H_{objective}$  (cf. Equation 2.9), on obtient :

$$\begin{aligned} W_{gdg'd'} &= -K_u [\delta_{gg'}(1 - \delta_{dd'}) + \delta_{gg'}(1 - \delta_{gg'})] \\ &\quad - K_o O_{gdg'd'}(1 - \delta_{gg'})(1 - \delta_{dd'}) \\ &\quad + K_c S_{gdg'd'}(1 - \delta_{gg'})(1 - \delta_{dd'}) \end{aligned} \quad (2.12)$$

$$I_{gd} = 2K_u \quad (2.13)$$

où  $\delta_{ij}$  est l'indice de Kroneker défini par :

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

Notons que la matrice des poids des connexions est symétrique, ce qui assure la convergence du processus de minimisation. On note aussi que les éléments diagonaux de cette matrice sont nuls.

A présent, nous avons un réseau de neurones de Hopfield défini par une matrice de poids et un vecteur d'entrées externes qui représentent les contraintes globales utilisées pour résoudre le problème de mise en correspondance.

#### 2.4.4.2 Processus de minimisation

Comme nous l'avons indiqué plus haut, nous utilisons un mode de fonctionnement continu pour la minimisation de la fonction d'énergie du réseau. Dans ce cas, l'état des neurones peut prendre n'importe quelle valeur de l'intervalle  $[0, 1]$ . Il est considéré comme une fonction continue de la variable temporelle qui est aussi continue. Dans le modèle continu, l'état  $E_{gd}$  d'un neurone  $n_{gd}$  représente une probabilité de vraisemblance ou une mesure de qualité d'appariement entre le point de contour  $g$  de l'image gauche et le point de contour  $d$  de l'image droite.

Le processus de minimisation consiste à faire évoluer le réseau vers un état stable correspondant à un minimum de la fonction d'énergie du réseau, et donc à un minimum de la fonction objective. Celui-ci est atteint lorsqu'aucun des neurones du réseau ne peut changer d'état. La stabilité du réseau est assurée du fait que la matrice des poids de connexions est symétrique [139].

L'évolution temporelle d'un neurone  $n_{gd}$  du réseau est régie par l'équation différentielle suivante [139] :

$$\frac{dU_{gd}}{dt} = J(U_{gd}) = -\frac{U_{gd}}{\tau} + \sum_{(g,d) \in \Omega} W_{gdg'd'} E_{gd} E_{g'd'} + I_{gd} \quad (2.14)$$

où  $U_{gd}$  est le potentiel d'entrée du neurone  $n_{gd}$ . Le paramètre  $\tau$  représente une constante de temps qui est ici identique pour tous les neurones du réseau. Le potentiel d'entrée  $U_{gd}$  et l'état  $E_{gd}$  du neurone  $n_{gd}$  sont liés par la fonction de transfert de type sigmoïde (cf. Figure 2.12) :

$$E_{gd} = g(U_{gd}) = \frac{1}{2} \left( 1 + \tanh \left( \frac{U_{gd}}{\lambda} \right) \right) \quad (2.15)$$

où  $\lambda$  est un paramètre qui est analogue à celui de la température pour un système physique. Il permet de contrôler l'étendue de la fonction de transfert entre les deux valeurs limites 0 et 1 (cf. Figure 2.12). On peut noter que lorsque  $\lambda$  prend des valeurs éloignées de 0, l'étendue est large. Inversement, lorsque  $\lambda$  tend vers 0, la fonction de transfert tend vers une fonction binaire, et par conséquent, le modèle continu du réseau de Hopfield tend vers le modèle binaire.

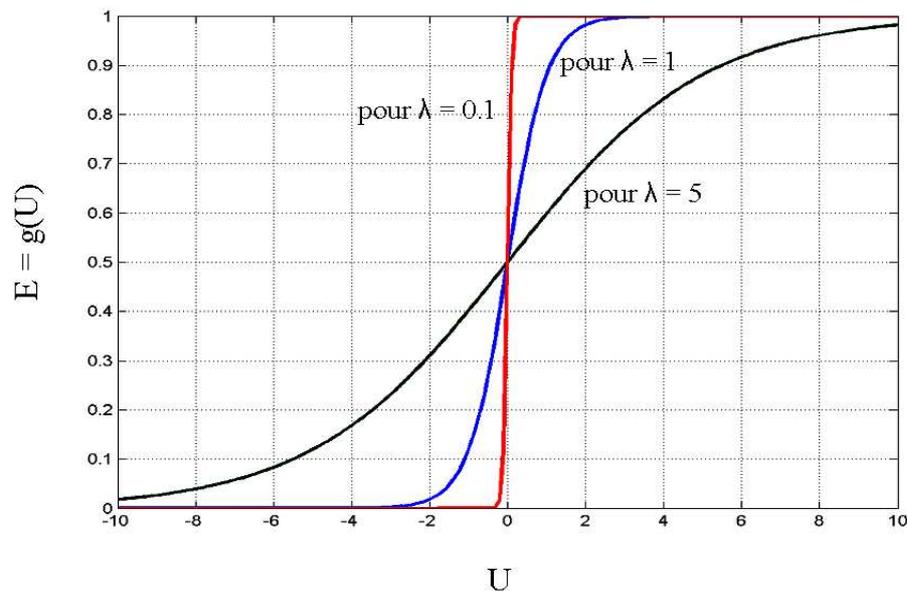


FIG. 2.12 – Fonction de transfert pour différentes valeurs de  $\lambda$ .

En partant d'un état initial, la procédure de mise à jour de l'état des neurones est lancée pour que le réseau atteigne un état stable. Les neurones sont consultés selon un schéma asynchrone séquentiel, c'est à dire à tour de rôle. Dès que l'état d'un neurone est actualisé, le nouvel état est pris en compte pour la mise à jour des autres neurones. La

procédure de mise à jour du neurone  $n_{gd}$  consiste tout d'abord à calculer son potentiel d'entrée  $U_{gd}$  à partir d'une résolution approchée de l'équation  $\frac{dU_{gd}}{dt} = J(U_{gd})$  (cf. Equation 2.14). On détermine ensuite le nouvel état du neurone  $n_{gd}$  en utilisant la fonction de transfert  $E_{gd} = g(U_{gd})$  (cf. Equation 2.15).

Il convient de noter que le processus de minimisation induit par le fonctionnement du réseau est équivalent à une descente dans un puits d'énergie. En d'autres termes, le réseau évolue vers un état stable correspondant à un minimum local de sa fonction d'énergie. Si l'état initial n'est pas trop loin de la bonne solution, cette descente permet de converger vers une solution acceptable en un temps considérablement réduit en comparaison avec le modèle stochastique qui réalise une recherche globale du minimum.

La stratégie d'initialisation que nous avons adoptée est très simple. Elle consiste à affecter une probabilité de vraisemblance égale à 0.5 à tous les couples candidats à l'appariement, c'est à dire :  $E_{gd} = 0.5$  quel que soit le couple  $(g, d)$  appartenant à l'ensemble  $\Omega$ . A partir de cet état initial, l'état des neurones qui correspondent à des appariements corrects (respectivement incorrects) tend vers 1 (respectivement 0) au cours de l'évolution du réseau.

### 2.4.5 Analyse de l'état final du réseau

Le résultat de mise en correspondance est donné par l'état final du réseau qui correspond à un minimum de sa fonction d'énergie, et donc de la fonction objective. Nous savons que l'état d'un neurone représente une probabilité de vraisemblance de l'appariement d'un point de contour de l'image gauche avec un point de contour de l'image droite. Nous savons également que chaque ligne (respectivement colonne) du réseau est constituée de l'ensemble des neurones correspondant aux appariements possibles d'un point de contour de l'image gauche (respectivement droite). L'extraction des paires de points de contour appariées consiste alors à chercher les neurones qui présentent les fortes réponses dans les lignes et les colonnes du réseau.

La recherche des neurones à réponse maximale se fait de la manière suivante. On cherche dans chaque ligne le neurone ayant la valeur de l'état la plus élevée. Cette pre-

mière recherche peut conduire à des situations où plusieurs neurones sont choisis dans une même colonne. Pour éviter ces situations qui correspondent à des appariements multiples, on effectue la même procédure de recherche dans chaque colonne du réseau.

D'autres stratégies peuvent être utilisées pour extraire les paires de points de contour appariées. On peut par exemple effectuer une recherche globale au sein du réseau en sélectionnant tout d'abord le neurone ayant l'état le plus élevé. On élimine ensuite les neurones situés dans la ligne et dans la colonne du neurone sélectionné. Puis, on réitère ces deux étapes jusqu'à ce que tous les neurones du réseau aient été visités. On peut aussi utiliser le même raisonnement, mais en imposant la contrainte d'ordre au cours de la recherche, c'est à dire en ne considérant que les neurones qui correspondent à des couples respectant la contrainte d'ordre avec les couples déjà extraits.

## 2.4.6 Choix des paramètres

Notre méthode de mise en correspondance utilise trois groupes de paramètres. Le premier est constitué des paramètres  $\theta$  et  $\alpha$  utilisés lors de l'évaluation de la contrainte de continuité de la disparité (cf. Equation 2.8). Le deuxième groupe de paramètres concerne les coefficients de pondération des termes de la fonction objective (cf. Equation 2.9). Le dernier groupe de paramètres est lié à la dynamique du réseau de Hopfield qui dépend des valeurs de  $\tau$  et  $\lambda$  (cf. Equation 2.15).

### 2.4.6.1 Choix des paramètres $\theta$ et $\alpha$

Les paramètres  $\theta$  et  $\alpha$  sont utilisés lors de l'évaluation de la compatibilité de deux couples de points de contour vis-à-vis de la contrainte de continuité de la disparité (cf. Equation 2.8).

Le paramètre  $\theta$  représente un facteur de tolérance. Il permet de prendre en compte les effets du bruit et les distorsions. Il est ajusté de telle sorte qu'une compatibilité maximale, c'est à dire proche de 1, soit obtenue pour les paires de couples ayant des disparités voisines, et qu'une compatibilité minimale, c'est à dire proche de  $-1$ , soit obtenue pour les paires de couples ayant des disparités éloignées.

Le paramètre  $\alpha$  contrôle la pente de la fonction de mesure de compatibilité en  $\theta$ . Plus  $\alpha$  est grand, plus la pente est forte, c'est à dire plus on tend vers une fonction binaire de valeurs  $-1$  et  $1$ . Inversement, plus  $\alpha$  est proche de  $0$ , plus la pente est faible, c'est à dire plus le passage entre  $-1$  et  $1$  est lisse.

Les tests expérimentaux ont montré qu'avec  $\theta = 20$  et  $\alpha = 0.1$ , on obtient des résultats satisfaisants.

#### 2.4.6.2 Choix des coefficients de pondération $K_u$ , $K_o$ et $K_c$

La stratégie adoptée pour choisir la valeur des coefficients de pondération des termes de la fonction objective (cf. Equation 2.9) est basée sur l'analyse de l'importance à donner à chacune des contraintes utilisées. Celles-ci n'ont pas obligatoirement la même importance vis-à-vis de la recherche des bons appariements.

Nous considérons que l'unicité est la contrainte devant avoir la plus forte contribution. En effet, c'est la contrainte la plus stable, contrairement aux contraintes d'ordre et de continuité de la disparité qui utilisent des hypothèses sur la scène observée. C'est d'ailleurs la contrainte la plus utilisée en stéréovision. Les nombreux tests effectués nous ont permis ainsi de choisir les valeur suivantes :

- Pour le terme associé à la contrainte d'unicité :  $K_u = 5$
- Pour le terme associé à la contrainte d'ordre :  $K_o = 1$
- Pour le terme associé à la contrainte de continuité :  $K_c = 1$

#### 2.4.6.3 Choix des paramètres $\tau$ et $\lambda$

Le paramètre  $\tau$  représente une constante de temps dans l'équation qui gouverne l'évolution temporelle de l'état des neurones (cf. Equation 2.14). Hopfield utilise une constante de temps identique et égale à l'unité pour tous les neurones du réseau [69]. D'autres auteurs proposent d'attribuer une grande valeur à  $\tau$  [50]. Nos test expérimentaux nous ont conduits à choisir une valeur égale à  $10$  pour  $\tau$ .

Le paramètre  $\lambda$  intervient dans la fonction de transfert utilisée pour calculer l'état d'un neurone à partir de son potentiel d'entrée (cf. Equation 2.15). Il contrôle l'étendue de cette fonction entre les deux valeurs limites  $0$  et  $1$  (cf. Figure 2.12). Plus  $\lambda$  prend des

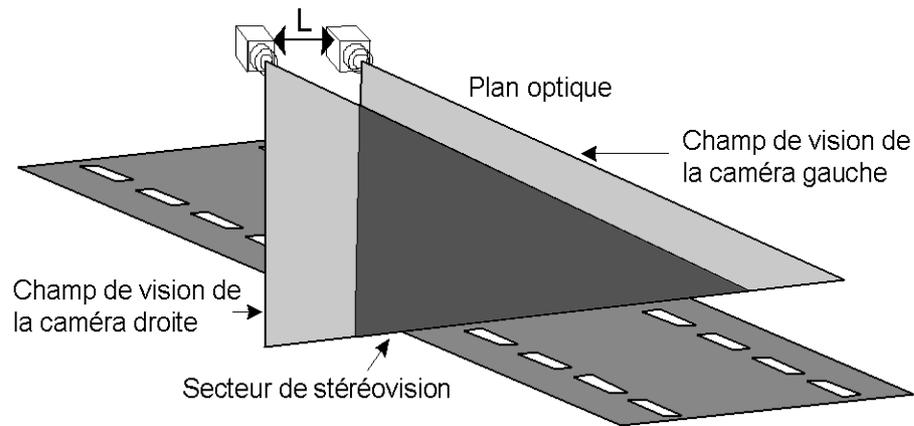
valeurs élevées, plus l'étendue est large. Inversement, plus  $\lambda$  est proche de 0, plus on tend vers une fonction de transfert binaire. Les tests que nous avons effectués nous ont permis de constater que lorsque  $\lambda$  est éloigné de 0, l'état des neurones prend des valeurs aux alentours de 0.5, ce qui conduit souvent à une oscillation du réseau sans convergence. Nous avons alors choisi une valeur très proche de 0 pour  $\lambda$ . Des résultats satisfaisants ont été obtenus avec  $\lambda = 0.01$ .

## 2.5 Résultats de mise en correspondance

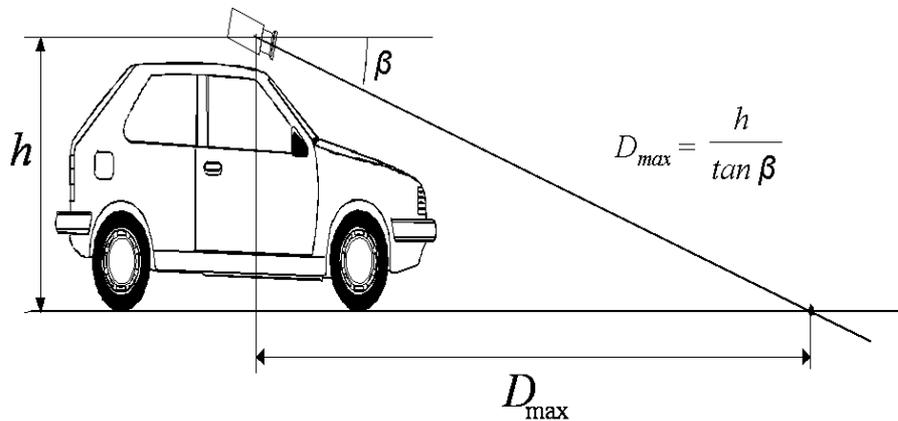
Quand j'ai rejoint l'équipe "Image et Décision", le programme PROMETHEUS était dans sa dernière année. Je n'ai pas pu alors intégrer mes procédures de mise en correspondance dans le démonstrateur mis en place pour évaluer les résultats obtenus à l'issue du programme. Par contre, mes prédécesseurs, J.-L. Bruyelle et J.-C. Burie, m'ont laissé de nombreuses séquences d'images stéréoscopiques acquises par le stéréoscope linéaire dans des conditions réelles de trafic.

Nous avons donc utilisé ces séquences pour évaluer les performances de notre méthode de mise en correspondance. Nous présentons ici quelques résultats issus de l'analyse de quelques séquences. Nous commençons tout d'abord par donner quelques précisions sur l'installation du stéréoscope linéaire sur le véhicule prototype.

Le stéréoscope linéaire est monté sur le toit d'un véhicule. Le plan de visée du stéréoscope est légèrement incliné d'un angle  $\beta$  par rapport à l'horizontale, de manière à couper la chaussée, qui est supposée plane, à une distance  $D_{max}$  à l'avant du véhicule (cf. Figure 2.13).



(a) Vue de dessus.



(b) Vue de côté.

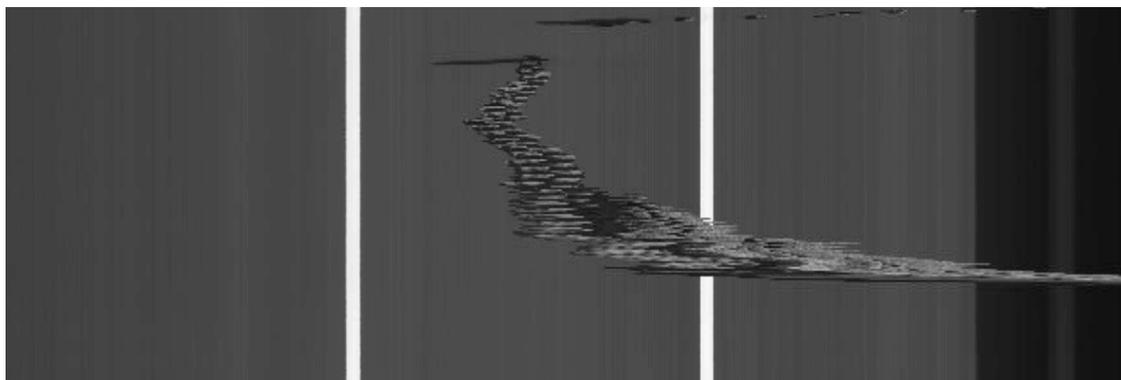
FIG. 2.13 – Stéréoscope installé sur le véhicule.

Les séquences d'images-ligne acquises sont représentées sous forme d'images bidimensionnelles en stockant les différentes images-ligne l'une après l'autre dans l'ordre de leur acquisition. Ces images bidimensionnelles s'interprètent de la manière suivante :

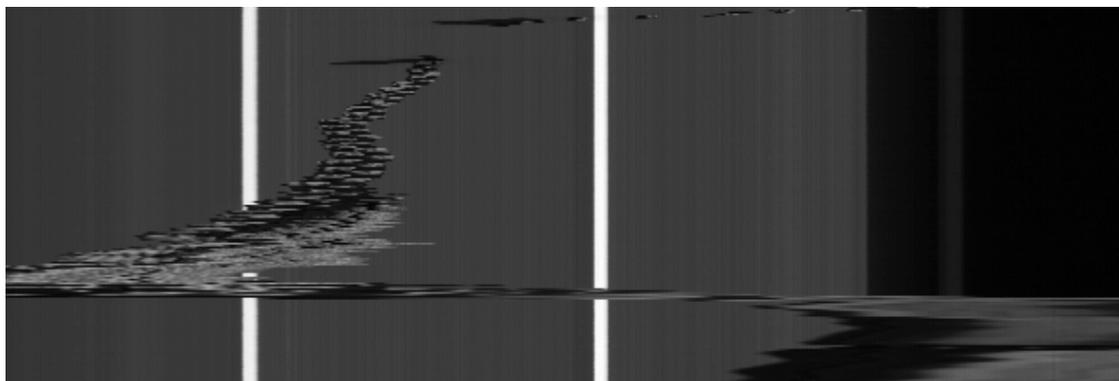
- L'axe horizontal correspond à la coupe que fait le plan de visée du stéréoscope avec la scène. Chaque ligne représente donc une image linéaire.
- L'axe vertical, orienté du haut vers le bas de l'image, représente l'axe du temps. La taille de chaque séquence est de 200 images-ligne.

### 2.5.1 Séquence 1 : Piéton évoluant devant le stéréoscope

Dans cette séquence stéréoscopique, la scène enregistrée correspond à un piéton qui se déplace devant le stéréoscope linéaire, monté sur le toit du véhicule et visant la chaussée à une distance de 50 mètres. La séquence stéréoscopique enregistrée est représentée sur la figure 2.14. Sur les images, on voit clairement les lignes blanches de la chaussée. Un véhicule situé hors du champ du stéréoscope projette une ombre, visible sur la droite des images comme une zone plus foncée que la chaussée.



(a) Séquence gauche.



(b) Séquence droite.

FIG. 2.14 – Séquence stéréoscopique "Piéton".

La figure 2.15 décrit schématiquement les déplacements du piéton au cours du temps. Partant du côté droit du stéréoscope (A), on le voit d'abord se diriger vers une zone située un peu au-delà de l'intersection du plan de visée et de la chaussée (B). Ce faisant, il passe hors de portée du stéréoscope linéaire et, en conséquence, disparaît des images. Quelques

secondes plus tard, il commence à se diriger vers le stéréoscope. Il réapparaît alors dans le champ de visée, en se dirigeant vers la caméra gauche (C), avant d'obliquer légèrement vers la caméra droite (D). Puis il part vers sa gauche et sort du secteur de stéréovision (E).

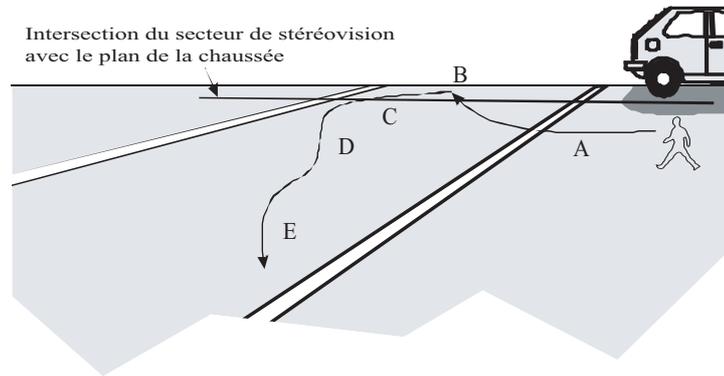


FIG. 2.15 – Description de la scène "Piéton".

L'analyse de cette séquence stéréoscopique par notre procédure de mise en correspondance consiste à traiter chaque couple d'images indépendamment des autres. A l'issue de chaque traitement, une liste des paires de points de contour appariées est constituée. Chaque paire de points de contour permet, grâce aux formules de triangulation géométrique, de déterminer la position d'un point de la scène réelle par rapport au référentiel lié au stéréoscope. Nous rappelons que la position d'un point de la scène réelle est définie dans le référentiel  $\{O, X, Z\}$  lié au stéréoscope par son abscisse  $X_P$  sur l'axe  $OX$  et sa distance  $Z_P$  sur l'axe  $OZ$  (cf. Section 2.2).

Tenant compte des paramètres intrinsèques des caméras, les formules de triangulation permettant de déterminer la position d'un point P de la scène dans le repère  $\{O, X, Z\}$  sont exprimées de la manière suivante :

$$Z_P = \frac{f \cdot L \cdot N}{l \cdot d} \quad (2.16)$$

$$X_P = \frac{l \cdot Z_P \cdot (x_g - c_g)}{f \cdot N} - \frac{L}{2} = \frac{l \cdot Z_P \cdot (x_d - c_d)}{f \cdot N} + \frac{L}{2} \quad (2.17)$$

où :

- $x_g$  et  $x_d$  constituent la paire de points de contour appariée.
- $c_g$  (respectivement  $c_d$ ) est la position du pixel central sur le capteur gauche (respectivement droite).
- $f$  est la distance focale des objectifs.
- $L$  est l'entre-axe du stéréoscope, c'est à dire la distance entre les centres optiques des capteurs gauche et droit.
- $N$  est la résolution du capteur, c'est à dire le nombre de pixels dans l'image.
- $l$  est la taille ou longueur des capteurs.

Les résultats issus du traitement de la séquence "Piéton" sont représentés sur l'image de reconstruction de la figure 2.16, où :

- L'axe vertical représente le temps. Celui-ci défilant du haut vers le bas de l'image.
- L'axe horizontal représente l'axe des abscisses  $OX$  du référentiel lié au stéréoscope. Un point se trouvant sur le bord gauche (respectivement droit) de l'image de reconstruction est situé à 5 mètres à gauche (respectivement droite) de l'axe optique du stéréoscope.
- La distance, exprimée sur l'axe  $OZ$  du référentiel lié au stéréoscope, est représentée par un niveau de gris. Celui-ci est d'autant plus clair que la distance par rapport au stéréoscope est grande.

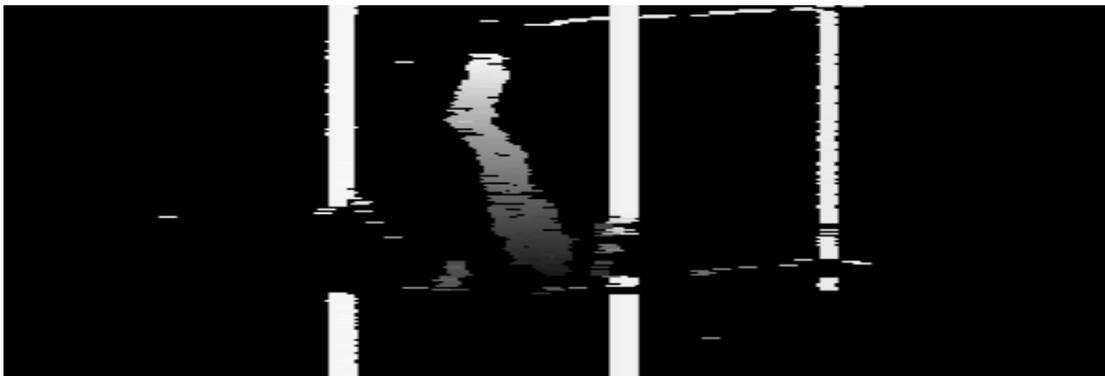


FIG. 2.16 – Image de reconstruction de la séquence "Piéton".

Les points de contour des bandes blanches sont correctement appariés et leur détection est

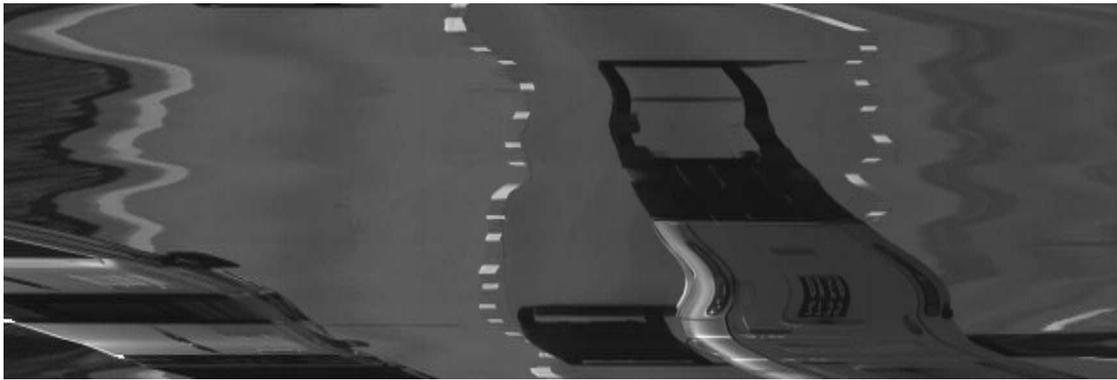
stable au cours du temps, c'est à dire que leurs positions ne varient pas. En effet, l'intersection du plan de visée et du plan de la chaussée ne change pas au cours de la séquence, puisque le stéréoscope est fixe. Les niveaux de gris clairs indiquent que les bandes blanches constituent les entités les plus éloignées du stéréoscope. La transition entre la chaussée et la zone d'ombre à droite de l'image de reconstruction est également bien détectée. La reconstruction des points de contour du piéton au cours du temps et leurs niveaux de gris, variant du clair au sombre, traduisent son rapprochement au stéréoscope suivant une trajectoire identique à celle produite dans la scène réelle.

On constate cependant la présence de faux appariements lorsque des occultations se produisent, c'est à dire lorsque le piéton cache l'une des bandes blanches à l'une des caméras. Ces faux appariements, qui sont peu nombreux, résultent de la mise en correspondance des points de contour d'une des bandes blanches avec ceux du piéton.

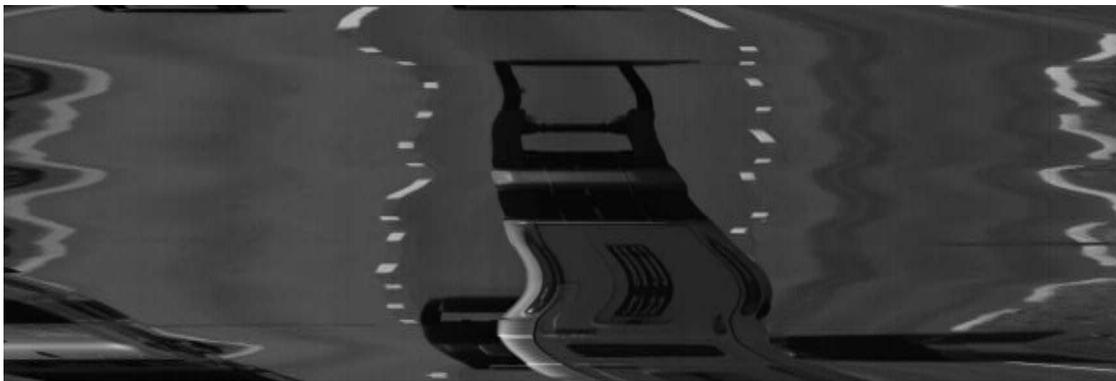
Au niveau des performances de rapidité, le traitement de la séquence stéréoscopique "Piéton" nécessite 2234 millisecondes environ sur un PC de type AMD Athlon XP 2800+ à 1.67 GHz avec 512 Mo de RAM. La séquence étant composée de 200 images-ligne, la cadence moyenne de traitement est d'environ 90 couples d'images-ligne par seconde.

### 2.5.2 Séquence 2 : Approche d'un véhicule

Dans la séquence stéréoscopique représentée sur la figure 2.17, le véhicule instrumenté est en circulation sur autoroute. Une description schématique de la scène enregistrée est donnée sur la figure 2.18.



(a) Séquence gauche.



(b) Séquence droite.

FIG. 2.17 – Séquence stéréoscopique "Approche d'un véhicule".

Le véhicule instrumenté roule sur la voie centrale derrière une voiture "Renault 4" (R4) en se rapprochant d'elle. Le plan de visée croise l'arrière de la R4 à une hauteur croissante à mesure que la distance décroît, ce qui explique l'inversion de l'image de la R4, dont on commence par voir les roues pour remonter jusqu'au toit (cf. Figure 2.17).

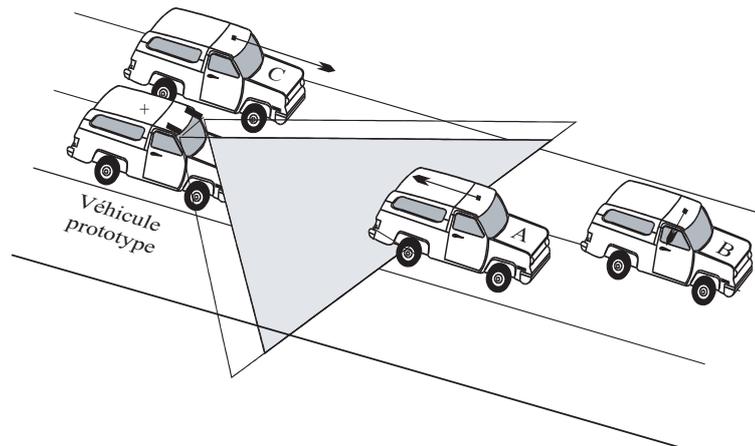


FIG. 2.18 – Description de la scène "Approche d'un véhicule". Les flèches indiquent les déplacements relatifs des véhicules par rapport au véhicule prototype.

Devant la R4 se trouve un autre véhicule, en train de se rabattre après l'avoir dépassée. Ce deuxième véhicule n'entre pas dans les champs de visée des deux caméras, mais son ombre est visible en bas au milieu et en bas à droite des images de la séquence (cf. Figure 2.17).

Par ailleurs, le véhicule prototype est lui-même dépassé par un troisième véhicule qui n'entre pas entièrement dans le champ de visée commun aux deux caméras. Ce véhicule apparaît en bas à gauche des images de la séquence (Cf. Figure 2.17). Enfin, on distingue clairement les bandes blanches de signalisation séparant la chaussée de l'autoroute en trois voies.

Appliquée à la séquence "Approche d'un véhicule", notre procédure de mise en correspondance fournit l'image de reconstruction de la figure 2.19. Un point se trouvant sur le bord gauche (respectivement droit) de l'image de reconstruction est situé à 7 mètres à gauche (respectivement droite) de l'axe optique du stéréoscope.

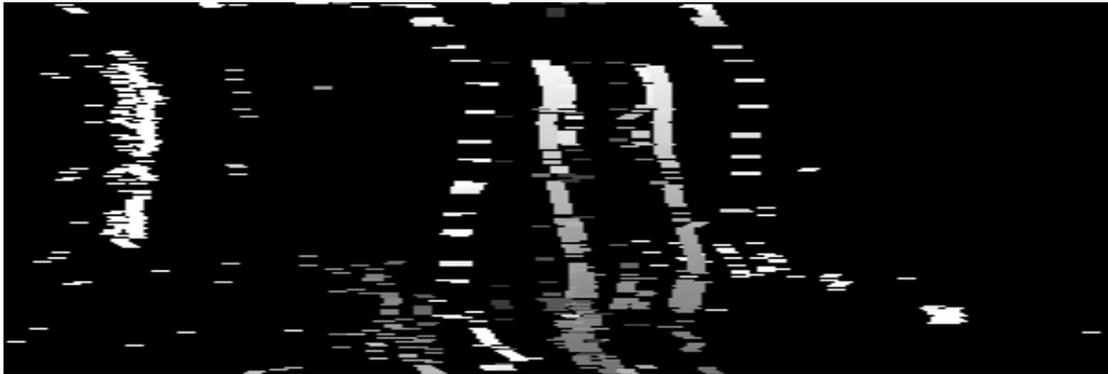


FIG. 2.19 – Image de reconstruction de la séquence "Approche d'un véhicule".

Sur l'image de reconstruction (cf. Figure 2.19), la plupart des points de contour de la R4 sont correctement reconstruits. On note que, contrairement à ce que l'on constatait sur les images brutes, la largeur apparente de la R4 est ici constante dans le temps, c'est à dire qu'elle reste identique pour toutes les lignes de l'image de reconstruction. Ceci est dû à la définition de l'axe horizontal du stéréoscope qui représente la largeur réelle des objets, et non pas celle de leur projection sur les capteurs. Les niveaux de gris des points de contour de la R4, de plus en plus foncés, traduisent son rapprochement au cours du temps par rapport au véhicule prototype.

Les deux bandes blanches en pointillés, qui séparent les voies de l'autoroute, sont également reconstruites. Leur ondulation est due au mouvement de lacet du véhicule instrumenté par rapport à la chaussée. En effet, le repère de travail est lié au véhicule portant le stéréoscope et non à la chaussée. La bande blanche gauche qui délimite le terre-plein central est également détectée, du moins lorsqu'elle est dans le champ de visée des deux caméras.

L'ombre projetée par le véhicule qui se rabat devant la R4 est clairement visible, en bas au milieu, sous la forme d'une ligne claire continue, et en bas à droite de l'image de reconstruction.

En bas à gauche de l'image de reconstruction, on distingue le véhicule qui double le véhicule instrumenté sous forme d'une ligne oblique très sombre en raison de sa proximité. On note quelques faux appariements dus à l'apparition momentanée, sur la séquence d'images issues de la caméra droite, de la bande blanche gauche qui délimite le terre-plein

central : l'algorithme de mise en correspondance semble confondre les points de contour de cette bande avec ceux du bord droit du véhicule extraits de l'image gauche.

Avec la même plate-forme de calcul, le traitement de la séquence stéréoscopique "Approche d'un véhicule" nécessite 781 millisecondes environ, ce qui correspond à une cadence moyenne de traitement d'environ 256 couples d'images-ligne par seconde.

## 2.6 Mise en correspondance par séparation des points de contour

Grâce à la puissance des machines de calcul de nos jours, la cadence de traitement de notre procédure de mise en correspondance est aujourd'hui compatible avec la détection d'obstacles temps réel. A l'époque, il a fallu utiliser une architecture spécialisée pour atteindre les performances requises pour une telle application. Cette limitation nous a poussé à chercher un autre moyen pour améliorer le temps des traitements. Nous avons ainsi proposé de mettre en correspondance les points de contour en les séparant en deux groupes : le groupe des points de contour dont le signe du gradient est positif et le groupe des points de contour dont le signe du gradient est négatif [142]. On décompose ainsi le problème de mise en correspondance en deux sous-problèmes identiques et indépendants. Chacun de ces sous-problèmes consiste à mettre en correspondance les points de contour de même nature. La mise en correspondance par séparation des points de contour nous a permis d'améliorer la cadence de traitement d'une manière significative. Cette technique permettrait d'augmenter davantage la cadence de traitement si on considérait une implantation parallèle de la mise en correspondance des deux groupes de points de contour. En revanche, nous avons constaté une perte au niveau de la qualité des résultats qui restent comparables à ceux présentés plus haut en terme de détection. Cette perte est due sans aucun doute au fait que les contraintes d'ordre et de continuité de la disparité, qui doivent être appliquées dans un contexte global, sont utilisées pour mettre en correspondance séparément les points de contour de chaque groupe.

## 2.7 Résultats au sein de l'équipe "Image et Décision"

Les premiers travaux de recherche que nous avons développés autour de l'approche neuronale ont joué un rôle moteur au sein de l'équipe "Image et Décision". En effet, la mise en correspondance d'images était une problématique prépondérante dans les travaux de l'équipe qui participait à de nombreux projets liés à la sécurité dans les transports terrestres. Suite à son évaluation sur les séquences d'images stéréoscopiques enregistrées dans le cadre du programme PROMETHEUS, l'approche neuronale que nous avons proposée s'est révélée très adaptée pour faire face à cette problématique.

Dans le cadre du projet "Carrefour Intelligent" de l'INRETS, l'équipe travaillait sur la localisation et le suivi de véhicules par des contours actifs extraits d'une séquence d'images stéréoscopiques. En collaboration avec C. Vieren, chercheur au sein de l'équipe, nous avons adapté et appliqué notre procédure pour mettre en correspondance les contours actifs modélisant des véhicules dans une séquence d'images stéréoscopiques [143]. La fonction objective est définie à partir des contraintes d'unicité, d'ordre et de continuité de la disparité. Les couples candidats à l'appariement sont déterminés par application de la contrainte épipolaire aux positions des barycentres des contours actifs : deux contours actifs constituent un couple candidat à l'appariement si leurs barycentres se trouvent à la même ordonnée, avec une certaine tolérance. Dans le cadre du même projet, P. Vannoorenberghe, ancien chercheur au sein de l'équipe "Image et Décision", actuellement Maître de Conférences à l'Université de Rouen, travaillait sur le suivi de piétons dans un environnement urbain par analyse de séquences d'images monoculaires [163]. Nous avons adapté et utilisé notre méthode neuronale pour mettre en correspondance les points d'intérêt extraits des images successives [164]. En plus de sa position dans l'image, un point d'intérêt est caractérisé par la moyenne des valeurs absolues des amplitudes des gradients directionnels. Les couples de points d'intérêt candidats à l'appariement sont déterminés par utilisation des contraintes de similarité et de continuité de mouvement. La contrainte de similarité suppose que deux points d'intérêt, extraits de deux images successives, engendrés par le même point de la scène, présentent des valeurs de gradient directionnel comparables. La fonction objective est construite à partir des contraintes d'unicité et

d'ordre.

Les résultats issus de ma thèse ont été également le point de départ d'autres travaux de recherche au sein de l'équipe "Image et Décision". Ces travaux auxquels j'ai participé ont fait l'objet de deux thèses réalisées au sein de l'équipe.

L'inconvénient majeur du dispositif de prise de vue linéaire est lié à l'impossibilité de modifier la distance de surveillance, prédéterminée lors du calibrage du stéréoscope. De plus, malgré la simplicité de la procédure de calibrage mise au point par J.-L. Bruyelle, assurer la coplanéité des plans de visée de deux caméras linéaires reste toujours une étape délicate à réaliser. Ces limitations ont poussé l'équipe à développer une nouvelle stratégie de détection d'obstacles grâce à un nouveau dispositif de prise de vue stéréoscopique, baptisé stéréoscope bimodal à portée variable. Composé de deux caméras matricielles, ce dispositif permet de simplifier avantageusement la procédure de calibrage, tout en disposant de plusieurs plans de surveillance : chaque couple de lignes des deux images correspond à une distance de surveillance. On peut alors adapter le choix des lignes traitées pour chercher simultanément les obstacles à différentes distances ou renforcer les traitements dans une zone où il y a des indices de présence d'un obstacle, tout en respectant la contrainte temps réel. Ce travail a fait l'objet de la thèse de D. Dooze qui a été réalisée dans le cadre du programme STATUE pour assurer la détection d'obstacles à l'avant des systèmes de transports guidés et automatisés [42]. J'ai participé à ce travail en m'occupant du problème de l'appariement des points de contour extraits des images issues du nouveau stéréoscope. L'approche neuronale que nous avons développée a été utilisée encore une fois de plus pour résoudre le problème. Les candidats à l'appariement sont obtenus grâce aux contraintes d'épipolarité, de position et d'orientation des points de contour. La fonction objective est définie pour évaluer les contraintes d'unicité, d'ordre et de continuité de la disparité.

Le processus de minimisation induit par un réseau de neurones de Hopfield n'est que local, c'est à dire qu'il fournit le premier minimum rencontré. Une autre orientation à nos travaux de recherche au sein de l'équipe "Image et Décision" a porté sur l'exploitation des algorithmes génétiques pour résoudre le problème de mise en correspondance. Cette orientation était motivée par le fait que ces algorithmes permettent une recherche globale

dans l'espace des solutions. Nous avons tout d'abord utilisé un algorithme génétique basé sur un codage binaire classique. Nous avons rapidement constaté son inefficacité tant au niveau de la qualité des appariements qu'au niveau du temps de convergence des algorithmes. Nous avons alors proposé un nouveau codage permettant à la fois une meilleure exploration de l'espace des solutions et une accélération des algorithmes. Ce travail a fait l'objet du DEA, puis de la thèse de H. Issa soutenue en 2004. Une synthèse de ce travail sera présenté au chapitre suivant.

## 2.8 Conclusion

Les travaux de recherche que nous avons présentés dans ce chapitre constituaient une poursuite des travaux qui ont été entrepris au sein de l'équipe "Image et Décision" dans le cadre du Programme PROMETHEUS pour assurer la détection d'obstacles à l'avant d'un véhicule routier par stéréovision linéaire. La problématique centrale concernait la mise en correspondance d'images issues d'un stéréoscope linéaire conçu et réalisé par J.-L. Bruyelle, ancien chercheur de l'équipe. Les méthodes de mise en correspondance qui ont été présentées par J.-C. Burie dans le cadre de sa thèse ayant montré leurs limites, j'ai proposé une nouvelle approche basée sur l'utilisation des réseaux de neurones de Hopfield. Le travail consistait à développer cette approche au sein de l'équipe qui participait à de nombreux projets liés à la sécurité dans les transports terrestres.

Cette nouvelle approche consiste à considérer le problème de mise en correspondance comme un problème d'optimisation où une fonction objective, définie à partir des contraintes du problème, doit être minimisée. Deux types de contraintes, locales et globales, sont utilisées. Les contraintes locales de position et de signe du gradient sont utilisées pour déterminer les couples candidats à l'appariement. Les contraintes globales d'unicité, d'ordre et de continuité de la disparité sont utilisées pour définir la fonction objective. Le processus de minimisation est implanté sous la forme d'un réseau de neurones de Hopfield dans lequel chaque neurone représente un couple de points de contour candidat à l'appariement.

Appliquée aux séquences d'images stéréoscopiques enregistrées dans le cadre du

---

programme PROMETHEUS, cette procédure de mise en correspondance a donné entière satisfaction tant au niveau de la qualité des appariements qu'au niveau de la cadence de traitement. Elle a été adaptée et utilisée avec succès au sein de l'équipe dans le cadre de nombreux projets pour mettre en correspondance d'autres types de primitives issues d'images stéréoscopiques et temporelles.

Ces premiers résultats ont donné lieu à d'autres travaux de recherche au sein de l'équipe "Image et Décision". Ces travaux auxquels j'ai participé ont fait l'objet de deux thèses réalisées au sein de l'équipe. La première concernait le développement d'une nouvelle stratégie de détection d'obstacle en utilisant un stéréoscope bimodal à portée variable. La deuxième thèse portait sur l'exploitation des algorithmes génétiques pour résoudre le problème de mise en correspondance. Une synthèse des travaux de cette thèse fera l'objet du chapitre suivant.

---

---

## CHAPITRE 3

---

Mise en correspondance par algorithmes  
génétiqnes

## 3.1 Introduction

L'approche neuronale que nous avons présentée dans le chapitre précédent a donné des résultats très satisfaisants tant au niveau de la qualité des mises en correspondance qu'au niveau de la cadence des traitements. Le processus de minimisation induit par un réseau de neurones de Hopfield étant local, nous avons orienté une partie de nos travaux de recherche vers l'exploitation des algorithmes génétiques qui sont des techniques de recherche globale. Dans un premier temps, nous présentons un algorithme génétique basé sur un codage binaire classique. Nous montrons ensuite les limites de ce codage. Afin d'améliorer les performances de notre algorithme génétique, nous proposons un nouveau codage permettant une meilleure exploration de l'espace des solutions. Ce travail a fait l'objet de la thèse de H. Issa, co-encadrée par J.-G. Postaire et moi-même. Cette thèse a été soutenue au sein de l'équipe "Scènes Dynamiques" en février 2004. Elle a donné lieu à une publication dans un journal, un chapitre de livre, quatre communications dans des conférences internationales et deux communications dans des conférences nationales.

## 3.2 Codage binaire

La première étape à considérer lors du développement d'un algorithme génétique est d'établir un codage permettant de représenter les solutions du problème sous forme de chromosomes constitués de gènes. Pour le problème de mise en correspondance, une représentation classique est d'utiliser des chromosomes où chaque gène représente une possibilité de mise en correspondance entre un point de contour de l'image gauche et un point de contour de l'image droite.

Soit  $G$  (respectivement  $D$ ) la liste des points de contour extraits de l'image gauche (respectivement droite). Soit  $N_G$  (respectivement  $N_D$ ) le nombre de points de contour dans la liste  $G$  (respectivement  $D$ ). Considérons la matrice de mise en correspondance  $\mathbf{E}$  de la figure 3.1, où chaque élément  $E_{gd}$  représente une possibilité de mise en correspondance entre les points de contour  $g$  de l'image gauche et  $d$  de l'image droite (cf. Chapitre 2). Seuls les éléments associés aux couples candidats à l'appariement, c'est à dire ceux qui

respectent les contraintes locales de position et de signe du gradient, sont considérés. Nous pouvons ainsi construire un chromosome sous la forme d'une chaîne binaire en parcourant la matrice de mise en correspondance, par exemple de bas en haut et de gauche à droite.

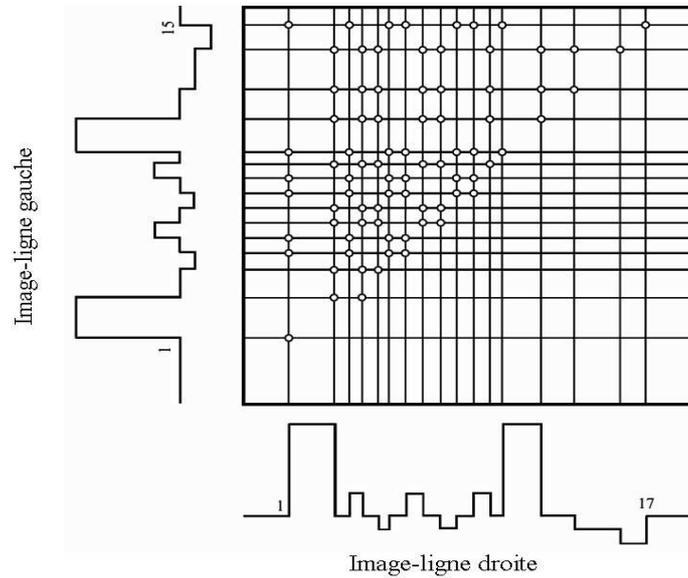


FIG. 3.1 – Exemple de matrice de mise en correspondance d'un couple d'images-ligne gauche et droite avec  $N_G = 15$  et  $N_D = 17$ . Seuls les couples qui vérifient les contraintes locales sont représentés.

Dans la pratique, un gène, noté  $((g, d), E_{gd})$ , représentant un couple  $(g, d)$  candidat à l'appariement est constitué de deux informations (cf. Figure 3.2). La première est statique et contient les indices  $g$  et  $d$  des deux points de contour gauche et droit dans leurs listes respectives. La deuxième information est dynamique et concerne la valeur, ou l'état,  $E_{gd}$  du gène. Celle-ci indique si le couple  $(g, d)$  est apparié ( $E_{gd} = 1$ ) ou non ( $E_{gd} = 0$ ).

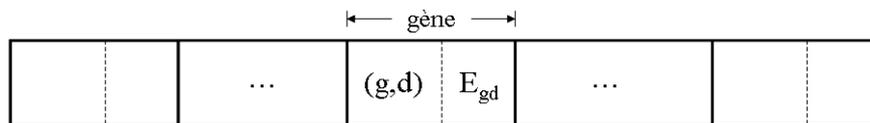


FIG. 3.2 – Chromosome binaire.

### 3.3 Evaluation des chromosomes binaires

Afin de permettre à une population de chromosomes d'évoluer vers une solution du problème, l'algorithme génétique a besoin d'une fonction, appelée fonction d'évaluation ou "fitness", pour évaluer l'adaptation, ou la qualité, de chaque chromosome. La fonction d'évaluation pour notre problème de mise en correspondance est définie à partir des contraintes globales d'unicité, d'ordre et de continuité de la disparité. Elle est construite de telle sorte que son minimum corresponde à une solution où les appariements respectent au mieux ces contraintes. Pour évaluer un chromosome binaire, nous nous servons de la matrice de mise en correspondance qui lui est associée. Ainsi, notre fonction d'évaluation n'est autre que la fonction objective qui a été introduite dans l'approche neuronale présentée dans le chapitre précédent. Elle est définie de la manière suivante :

$$\begin{aligned}
H_{binaire} = & K_u \left[ \sum_{g \in G} \left( 1 - \sum_{d \in D / (g,d) \in \Omega} E_{gd} \right)^2 + \sum_{d \in D} \left( 1 - \sum_{g \in G / (g,d) \in \Omega} E_{gd} \right)^2 \right] \\
& + K_m \left( N_{min} - \sum_{g \in G} \sum_{d \in D / (g,d) \in \Omega} E_{gd} \right)^2 \\
& + K_o \sum_{(g,d) \in \Omega} \sum_{\substack{(g',d') \in \Omega \\ g' \neq g \text{ et } d' \neq d}} O_{gdg'd'} E_{gd} E_{g'd'} \\
& - K_c \sum_{(g,d) \in \Omega} \sum_{\substack{(g',d') \in \Omega \\ g' \neq g \text{ et } d' \neq d}} S_{gdg'd'} E_{gd} E_{g'd'} \tag{3.1}
\end{aligned}$$

où  $N_{min} = \min(N_G, N_D)$  et où  $K_u$ ,  $K_m$ ,  $K_o$  et  $K_c$  sont des coefficients positifs permettant de pondérer les différents termes de la fonction d'évaluation.  $\Omega$  représente l'ensemble des couples candidats à l'appariement, c'est à dire les couples de points de contour qui respectent les contraintes locales de position et de signe du gradient :

$$\Omega = \{(g, d) \in G \times D / (g,d) \text{ vérifie les contraintes de position et de signe du gradient}\}$$

Le premier terme de la fonction d'évaluation sert à défavoriser les solutions qui mettent en défaut la contrainte d'unicité, c'est à dire favoriser les solutions qui mettent

en évidence un seul appariement dans chaque ligne et dans chaque colonne de la matrice de mise en correspondance. Le second terme est introduit pour renforcer les solutions mettant en évidence un maximum de couples appariés. Ce terme est en accord avec la contrainte d'unicité, puisque le nombre maximal de couples appariés est limité à  $N_{min} = \min(N_G, N_D)$ . Le troisième terme représente la contrainte d'ordre. Le coefficient  $O_{gdg'd'}$  représente une pénalité qui intervient au cas où la contrainte n'est pas respectée par les couples  $(g, d)$  et  $(g', d')$ . Il s'exprime de la manière suivante :

$$O_{gdg'd'} = |\text{signe}(x_g - x_{g'}) - \text{signe}(x_d - x_{d'})|$$

$x_g, x_{g'}, x_d$  et  $x_{d'}$  sont respectivement les positions des points de contour  $g, g', d$  et  $d'$ . La fonction signe est définie comme suit :

$$\text{signe}(a) = \begin{cases} 1 & \text{si } a > 0 \\ 0 & \text{sinon} \end{cases}$$

Le quatrième terme est utilisé pour renforcer la contrainte de continuité de la disparité. Le coefficient  $C_{gdg'd'}$  représente une mesure de compatibilité entre les couples candidats  $(g, d)$  et  $(g', d')$  vis-à-vis de cette contrainte. Il s'exprime de la façon suivante (cf. Chapitre 2) :

$$S_{gdg'd'} = S(X_{gdg'd'}) = \frac{2}{1 + e^{\alpha(X_{gdg'd'} - \theta)}} - 1 \quad (3.2)$$

où  $X_{gdg'd'}$  est la valeur absolue de la différence des disparités des couples  $(g, d)$  et  $(g', d')$  :

$$X_{gdg'd'} = |(x_g - x_d) - (x_{g'} - x_{d'})|$$

La mesure de compatibilité  $S(X_{gdg'd'})$  varie de manière continue entre  $-1$  et  $1$ . Elle est d'autant plus forte, c'est à dire proche de  $1$ , que les disparités des couples  $(g, d)$  et  $(g', d')$  sont voisines. Inversement, elle est d'autant plus faible, c'est à dire proche de  $-1$ , que la différence entre les disparités des couples  $(g, d)$  et  $(g', d')$  est importante.

Nous rappelons que le paramètre  $\theta$  contrôlant la position où la fonction  $S(X_{gdg'd'})$

s'annule, permet une certaine tolérance afin de prendre en compte les effets du bruit et les distorsions. Le paramètre  $\alpha$  contrôle la pente de cette fonction  $S(X_{gdg'd'})$  en  $\theta$ .

### 3.4 Mise en oeuvre de l'algorithme génétique

Une fois que le codage sous forme de chromosomes et que la fonction d'évaluation sont établis, l'étape suivante consiste à mettre en oeuvre un algorithme génétique pour construire un chromosome ayant une meilleure valeur d'adaptation vis-à-vis de la fonction d'évaluation. Dans notre cas, plus la valeur d'adaptation est petite, plus on se rapproche du meilleur chromosome, c'est à dire du chromosome qui minimise la fonction d'évaluation. Notre algorithme génétique, représenté sur l'organigramme de la figure 3.3, se déroule de la manière suivante :

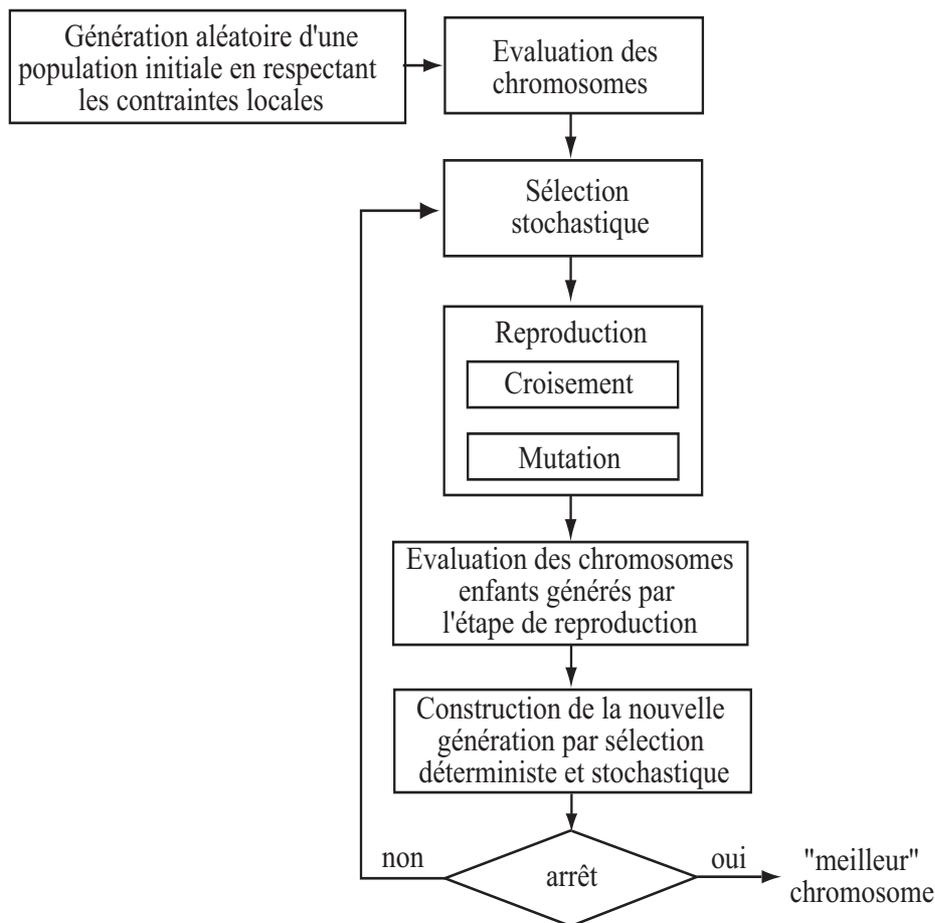


FIG. 3.3 – Schéma de l'algorithme génétique.

- **Etape 1** : Elle consiste à créer aléatoirement une population initiale de chromosomes. Nous rappelons que chacun de ces chromosomes ne doit coder que les couples de points de contour vérifiant les contraintes locales de position et de signe du gradient. Le nombre de chromosomes de la population  $Taille_{pop}$  demeure constant tout au long de l'évolution de la population.
- **Etape 2** : Elle consiste à évaluer les chromosomes de la population initiale en utilisant la fonction d'évaluation (cf. Equation 3.1).
- **Etape 3** : A partir de la population courante, on procède à la sélection des chromosomes parents qui participeront à la phase de reproduction. La procédure de sélection est basée sur le principe de la roulette de casino. Chaque chromosome occupe un secteur d'une roulette dont l'angle est proportionnel à sa valeur d'adaptation, fournie par la fonction d'évaluation. Pour sélectionner les chromosomes parents, on lance la roulette autant de fois qu'il y a de chromosomes dans la population. Ainsi, certains chromosomes peuvent être sélectionnés plusieurs fois et d'autres peuvent ne pas être sélectionnés. Les chromosomes sélectionnés constituent ce que nous appelons une population intermédiaire.
- **Etape 4** : Elle concerne la reproduction des chromosomes parents par croisement. En prenant les chromosomes de la population intermédiaire dans l'ordre de leur sélection, des paires de chromosomes sont tout d'abord constituées. Il s'agit ensuite, pour chaque paire de chromosomes, de tirer une valeur aléatoire comprise entre 0 et 1. Si cette valeur est inférieure ou égale à la probabilité de croisement  $P_c$ , alors les deux chromosomes sont croisés et les chromosomes enfants les remplacent dans la population intermédiaire. Sinon, les chromosomes parents sont maintenus. L'opérateur de croisement en un point est utilisé en raison de la simplicité de sa mise en oeuvre.
- **Etape 5** : On procède à la mutation des chromosomes de la population intermédiaire issue de l'étape de reproduction (étape 4). La procédure de mutation est appliquée à chaque chromosome en parcourant tous ses gènes. Pour chaque gène, on tire une valeur aléatoire comprise entre 0 et 1. Si cette valeur est inférieure ou égale à la probabilité de mutation  $P_m$ , on mute le gène. La mutation d'un

gène  $((g, d), E_{gd})$  ne concerne que l'information dynamique  $E_{gd}$  qui code l'état de mise en correspondance du couple de points de contour  $(g, d)$  (cf. Section 3.2). La mutation consiste à remplacer l'état du gène par son complément, c'est à dire :

$$E_{gd} = \begin{cases} 1 & \text{si } E_{gd} = 0 \\ 0 & \text{si } E_{gd} = 1 \end{cases}$$

- **Etape 6** : On évalue les chromosomes de la population intermédiaire issue de la phase de mutation (étape 5).
- **Etape 7** : Il s'agit de construire la nouvelle population. Pour cela, on fusionne les populations courante et intermédiaire et on effectue une sélection déterministe suivie d'une sélection stochastique. La sélection déterministe, basée sur une stratégie d'élitisme, permet de récupérer les "meilleurs" chromosomes présents dans les populations courante et intermédiaire. Les meilleurs chromosomes sont ceux qui présentent les meilleures valeurs d'adaptation. Le nombre de chromosomes fournis par la sélection déterministe représente une proportion de 10% de la taille de la population  $Taille_{pop}$ . Pour compléter la nouvelle population, la sélection stochastique, identique à celle de la roulette de casino décrite à l'étape 3, est appliquée aux chromosomes restants, c'est à dire à l'ensemble des chromosomes des populations courante et intermédiaire démunies des "meilleurs" chromosomes, obtenus par la sélection déterministe. Les chromosomes issus de ces deux sélections constituent alors la nouvelle population pour la génération suivante.
- **Etape 8** : Elle concerne le test d'arrêt de l'algorithme génétique. Nous avons choisi d'arrêter l'évolution de la population à l'issue d'un certain nombre de générations  $Nb_{gen}$ , fixé préalablement. Si la condition d'arrêt est atteinte, on récupère le "meilleur" chromosome de la population finale, sinon, on répète les étapes 3 à 8.

A l'issue de l'algorithme génétique, les résultats de mise en correspondance sont extraits du "meilleur" chromosome de la population finale, où les couples appariés sont ceux qui correspondent aux gènes ayant un état égal à 1.

Comme il manipule des chromosomes binaires, l'algorithme génétique présenté ci-dessus est baptisé algorithme génétique binaire (AGB).

## 3.5 Résultats de mise en correspondance issus de l'AGB

L'algorithme génétique décrit dans la section précédente a été appliqué aux séquences d'images stéréoscopiques "Piéton" et "Approche d'un véhicule", présentées dans le chapitre 2. Nous commençons tout d'abord par présenter les valeurs que nous avons choisies pour les différents paramètres utilisés. Ces valeurs sont choisies à la suite de tests expérimentaux.

### 3.5.1 Choix des paramètres

#### 3.5.1.1 Paramètres intervenant dans la fonction d'évaluation

Six paramètres sont utilisés dans la fonction d'évaluation. Quatre d'entre eux représentent les coefficients de pondération des termes de la fonction d'évaluation. Ils sont choisis de la manière suivante :

- Le coefficient de pondération du terme associé à la contrainte d'unicité :  $K_u = 5$ .
- Le coefficient de pondération du terme favorisant le maximum de couples appariés :  $K_m = 5$ .
- Le coefficient de pondération du terme associé à la contrainte d'ordre :  $K_o = 5$ .
- Le coefficient de pondération du terme associé à la contrainte de continuité de la disparité :  $K_c = 1$ .

Les deux autres paramètres interviennent dans la fonction de mesure de compatibilité lors de l'évaluation de la contrainte de continuité de la disparité (cf. Equation 3.2). Ils sont choisis comme suit :

- Le paramètre de tolérance  $\theta$  est fixé à 20.
- Le paramètre  $\alpha$  qui contrôle la pente de la fonction de mesure de compatibilité en  $\theta$  est fixé à 1.

### 3.5.1.2 Paramètres intervenant dans l'algorithme génétique

Notre algorithme génétique fait intervenir quatre paramètres :

- La probabilité de croisement  $P_c$  : Ce paramètre intervient pour décider du croisement d'une paire de chromosomes parents. Généralement, ce paramètre est choisi entre 0.5 et 0.9 [112]. Les tests effectués nous ont permis de fixer la probabilité de croisement à 0.6.
- La probabilité de mutation  $P_m$  : La taille d'un chromosome, c'est à dire le nombre de ses gènes, varie d'un couple d'images à l'autre. Elle est égale au nombre de couples de points de contour de l'ensemble  $\Omega$ , c'est à dire l'ensemble des couples candidats à l'appariement, obtenus après application des contraintes locales de position et de signe du gradient (cf. Section 3.3). Pour que l'opération de mutation des chromosomes ne perturbe pas le fonctionnement de l'algorithme génétique d'un couple d'images à un autre, nous avons choisi d'adapter la probabilité de mutation en prenant en compte la taille des chromosomes. Ainsi, pour chaque couple d'images à apparier, nous utilisons une probabilité de mutation  $P_m = \frac{1}{Taille_{chrom}}$ , où  $Taille_{chrom}$  est la taille du chromosome.
- Les deux derniers paramètres sont la taille de la population  $Taille_{pop}$ , c'est à dire le nombre de chromosomes dans la population, et le nombre de générations  $Nb_{gen}$ , c'est à dire le nombre d'itérations de l'algorithme génétique. Des résultats satisfaisants ont été obtenus avec  $Taille_{pop} = 100$  et  $Nb_{gen} = 300$ . Cependant, nous avons remarqué que ces deux paramètres doivent être choisis en fonction de la combinatoire de la mise en correspondance qui dépend de la taille des listes de points de contour extraits de chaque image. Plus la combinatoire est importante, plus il est nécessaire d'augmenter la taille de la population et le nombre de générations. Par conséquent, le temps de convergence, c'est à dire le temps de calcul nécessaire à l'algorithme génétique pour parvenir à la population finale, devient grand. On peut dire alors que ces deux paramètres contrôlent le temps de convergence de l'algorithme génétique. Une stratégie de réglage de ces deux paramètres est présentée plus loin, dans la section 3.9.

### 3.5.2 Résultats de mise en correspondance

Les figures 3.4 et 3.5 représentent respectivement les images de reconstruction obtenues après application de l'AGB aux séquences stéréoscopiques "Piéton" et "Approche d'un véhicule".

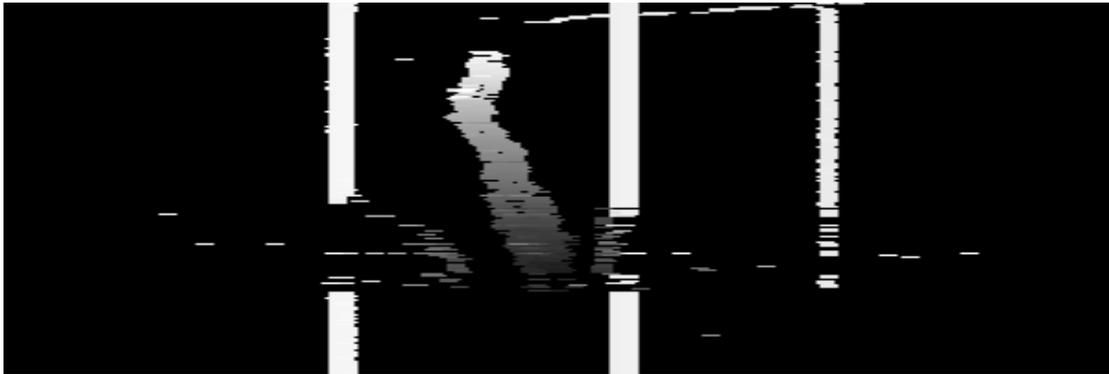


FIG. 3.4 – Séquence "Piéton" : Image de reconstruction obtenue par l'AGB.



FIG. 3.5 – Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par l'AGB.

Les résultats sont globalement similaires à ceux obtenus grâce à la méthode de mise en correspondance par réseau de neurones de Hopfield, décrite au chapitre 2 (cf. Figures 3.6 et 3.7). En revanche, la méthode par réseau de neurones de Hopfield est plus rapide (cf. Table 3.1).

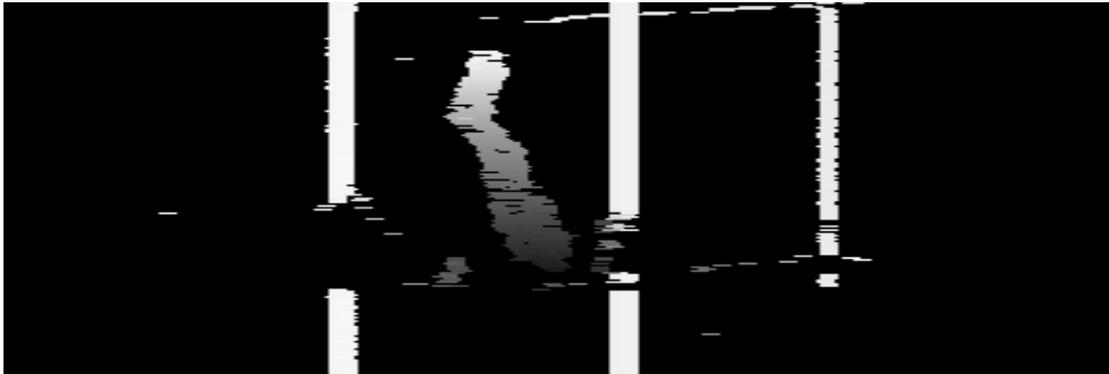


FIG. 3.6 – Séquence "Piéton" : Image de reconstruction obtenue par l'approche neuronale.

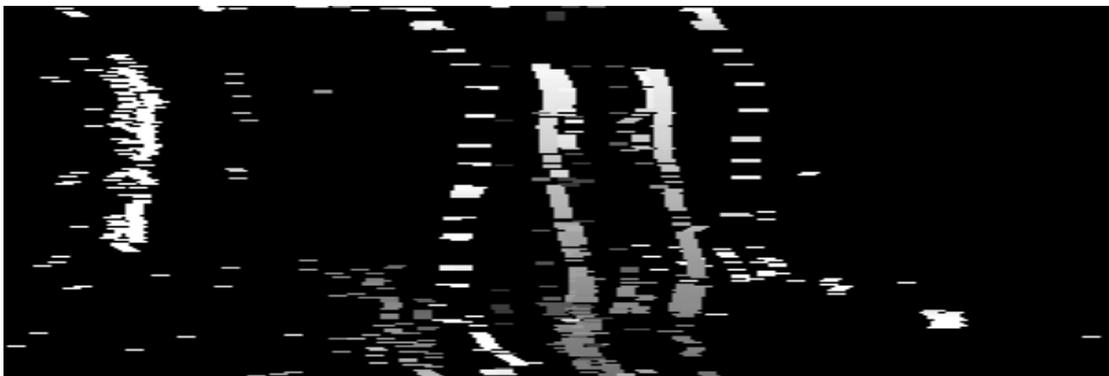


FIG. 3.7 – Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par l'approche neuronale.

TAB. 3.1 – Comparaison des cadences de traitement de l'AGB et de la méthode neuronale.

Méthode	Cadence de traitement (en nombre de couples d'images-ligne par seconde)	
	Séquence "Piéton"	Séquence "Approche d'un véhicule"
AGB	0.1	0.3
Méthode neuronale	90	256

### 3.5.3 Conclusion

Dans cette première partie, nous avons présenté un algorithme génétique permettant de réaliser la mise en correspondance des points de contours extraits d'images sté-

réoscopiques. Les solutions de mise en correspondance étant codées par des chromosomes binaires, l'algorithme génétique cherche à construire un chromosome ayant la meilleure valeur d'adaptation possible vis-à-vis de la fonction d'évaluation. Celle-ci est définie de telle sorte que sa valeur minimale corresponde à une solution où les appariements respectent au mieux les contraintes imposées.

Les résultats de mise en correspondance issus de l'AGB sont comparables à ceux obtenus par la méthode neuronale. Cependant, le temps de calcul nécessaire à l'AGB le rend incompatible avec une application temps réel telle que la détection d'obstacle à l'avant d'un véhicule. Cette limitation nous a poussés à analyser cet algorithme en termes de codage afin d'améliorer ses performances tant au niveau de la qualité des résultats qu'au niveau du temps des traitements.

## 3.6 Limitations du codage binaire

Reprenons l'exemple de la matrice de mise en correspondance présentée dans la section 3.2 (cf. Figure 3.8). Nous savons que le codage binaire et la notion de matrice de mise en correspondance sont équivalents. En effet, un chromosome binaire n'est autre que la chaîne composée des éléments de la matrice de mise en correspondance qui lui est associée (cf. Section 3.2).

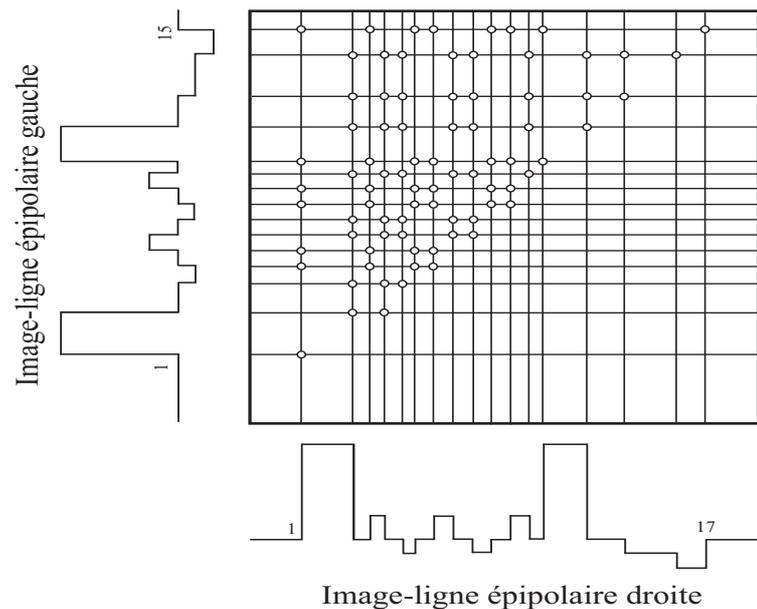


FIG. 3.8 – Exemple de matrice de mise en correspondance d'un couple d'images-ligne gauche et droite avec  $N_G = 15$  et  $N_D = 17$ . Seuls les couples qui respectent les contraintes locales sont représentés.

Une matrice de mise en correspondance peut mettre en évidence plusieurs appariements multiples à la fois sur ses lignes et sur ses colonnes (cf. Figure 3.8). On parle alors de matrice "très ambiguë", et par équivalence avec le codage binaire, on parle de chromosome "très ambigu". La présence d'un ou plusieurs chromosomes "très ambiguës" dans une population peut conduire l'évolution de l'algorithme génétique vers une mauvaise solution. Ces chromosomes peuvent apparaître soit à l'initialisation dans la population initiale, soit au cours de l'évolution de la population suite aux opérations de croisement et de mutation. Pour pouvoir arriver à une solution acceptable, le seul moyen est d'augmenter le nombre de générations, voire même augmenter la taille de la population.

On constate donc que le codage binaire ne permet pas une exploration efficace de l'espace des solutions. Par ailleurs, le codage binaire produit des chromosomes de grande taille, ce qui nécessite beaucoup de temps de calcul.

Pour pallier les limitations du codage binaire, nous proposons dans la section suivante un nouveau codage qui fournit des chromosomes compacts avec moins d'ambiguïtés vis-à-vis de la contrainte d'unicité. Nous montrons par la suite que ce nouveau codage

permet une meilleure exploration de l'espace des solutions tant au niveau de la qualité du résultat qu'au niveau des ressources en termes du nombre de générations et de la taille de la population.

## 3.7 Nouveau codage : le codage entier

### 3.7.1 Construction du nouveau codage

Au lieu de coder les états de mise en correspondance des couples de points de contour candidats à l'appariement, l'idée est de construire un chromosome ne mettant en évidence que les couples appariés, c'est à dire ceux ayant un état de mise en correspondance égal à 1.

Soient :

- $G$  (resp.  $D$ ) la liste des points de contour extraits de l'image gauche (resp. droite).  
On suppose que dans ces deux listes, les points de contour apparaissent dans le même ordre que celui de leur apparition dans les images brutes.
- $N_G$  (resp.  $N_D$ ) le nombre de points de contour dans  $G$  (resp.  $D$ ).
- $N_{min} = \min(N_G, N_D)$  et  $N_{max} = \max(N_G, N_D)$ .
- $T_{max} = \{1, 2, \dots, N_{max}\}$  la liste ordonnée des indices allant de 1 à  $N_{max}$ . Elle représente la liste des points de contour ayant la plus grande taille, c'est à dire  $N_{max}$ . De la même façon,  $T_{min} = \{1, 2, \dots, N_{min}\}$  est la liste ordonnée des indices allant de 1 à  $N_{min}$ . Elle représente la liste des points de contour ayant la plus petite taille, c'est à dire  $N_{min}$ . En d'autres termes, si  $N_G = \max(N_G, N_D)$ , alors  $T_{max}$  (resp.  $T_{min}$ ) représente la liste  $G$  (resp.  $D$ ). Inversement, si  $N_D = \max(N_G, N_D)$ , alors  $T_{max}$  (resp.  $T_{min}$ ) représente la liste  $D$  (resp.  $G$ ).

Le nouveau codage, baptisé codage entier, consiste en une chaîne  $C$  indexée par les éléments de la liste  $T_{max}$  et dont les valeurs appartiennent à la liste  $T_{min} \cup \{0\}$  (cf. Figure 3.9). Un chromosome entier est donc composé de  $N_{max}$  gènes. La valeur  $C_i$  du gène  $i$ ,  $i \in T_{max}$ , est un élément de la liste  $T_{min} \cup \{0\}$ . En d'autres termes, si  $C_i = 0$  alors le point de contour  $i$  de la liste  $T_{max}$  n'a pas de correspondant. Sinon, le couple de points

de contour  $(i, C_i)$  est apparié. Nous rappelons que, comme pour le codage binaire, seuls les couples de points de contour respectant les contraintes locales de position et de signe du gradient sont pris en compte.



FIG. 3.9 – Chromosome entier.

Pour bien comprendre le codage entier, prenons l'exemple de la matrice de mise en correspondance de la figure 3.10, où les listes des points de contour  $G$  et  $D$  sont de tailles  $N_G = 6$  et  $N_D = 5$ , respectivement. Dans ce cas, les listes  $T_{max} = \{1, 2, 3, 4, 5, 6\}$  et  $T_{min} = \{1, 2, 3, 4, 5\}$  représentent les listes  $G$  et  $D$ , respectivement. Le chromosome de la figure 3.11 se traduit alors de la manière suivante :

- $C_1 = 1$  signifie que le couple de points de contour  $(1, 1)$  est apparié.
- $C_2 = 2$  signifie que le couple de points de contour  $(2, 2)$  est apparié.
- $C_3 = 2$  signifie que le couple de points de contour  $(3, 2)$  est apparié.
- $C_4 = 3$  signifie que le couple de points de contour  $(4, 3)$  est apparié.
- $C_5 = 0$  signifie que le point de contour gauche 5 n'a pas de correspondant.
- $C_6 = 4$  signifie que le couple de points de contour  $(6, 4)$  est apparié.

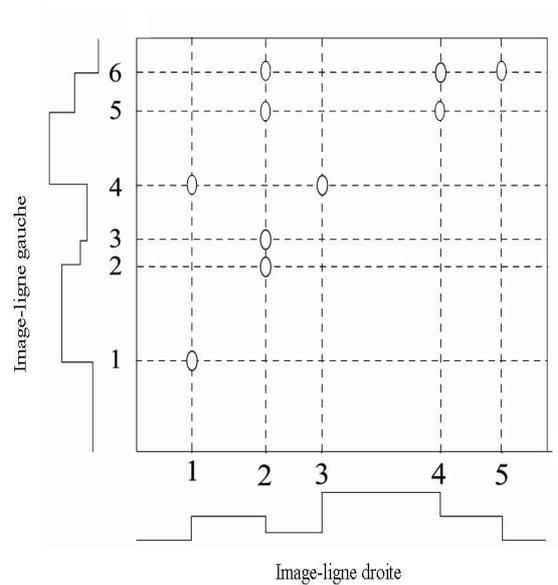


FIG. 3.10 – Exemple de matrice de mise en correspondance d'un couple d'images-ligne gauche et droite avec  $N_G = 6$  et  $N_D = 5$ . Seuls les couples vérifiant les contraintes locales sont représentés.

C:	$C_1=1$	$C_2=2$	$C_3=2$	$C_4=3$	$C_5=0$	$C_6=4$
$i=$	1	2	3	4	5	6

FIG. 3.11 – Exemple d'un chromosome généré par le codage entier :  $i \in T_{max} = \{1, 2, 3, 4, 5, 6\}$  et  $C_i \in \{0\} \cup T_{min}$  avec  $T_{min} = \{1, 2, 3, 4, 5\}$ .

L'opération de mutation définie pour le codage binaire n'est pas valable pour le codage entier. Il faut donc introduire une nouvelle opération pour effectuer la mutation des gènes des chromosomes entiers. On sait que les valeurs des gènes d'un chromosome entier appartiennent à  $T_{min} \cup \{0\}$ . Ainsi, la mutation du gène numéro  $i$  d'un chromosome entier  $C$  consiste à remplacer sa valeur  $C_i$  par une valeur tirée aléatoirement dans  $T_{min} \cup \{0\} - \{C_i\}$ .

Notons qu'il est possible de passer du codage entier au codage binaire. Supposons que les listes  $T_{max}$  et  $T_{min}$  représentent la liste des points de contour gauches  $G$  et la liste des points de contour droits  $D$ , respectivement. Dans ce cas, quel que soit le couple de

points de contour  $(g, d)$  appartenant à  $\Omega$ , ensemble des couples candidats à l'appariement (cf. Section 3.3), si  $C_g = d$ , alors  $E_{gd} = 1$ , sinon,  $E_{gd} = 0$  (cf. Figure 3.12). Dans le cas inverse, c'est à dire si on suppose que les listes  $T_{max}$  et  $T_{min}$  représentent la liste des points de contour droits  $D$  et la liste des points de contour gauches  $G$ , respectivement, quel que soit le couple de points de contour  $(g, d)$  appartenant à  $\Omega$ , si  $C_d = g$ , alors  $E_{gd} = 1$ , sinon,  $E_{gd} = 0$ .

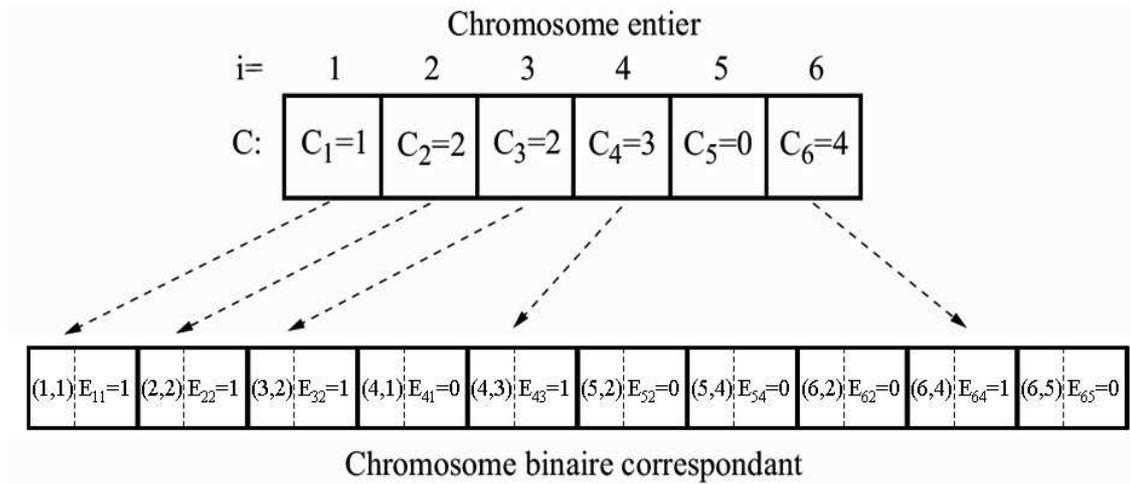


FIG. 3.12 – Chromosome entier et le chromosome binaire correspondant.

### 3.7.2 Intérêt du nouveau codage

Dans cette section, nous allons comparer les codages binaire et entier et montrer que ce dernier permet une meilleure exploration de l'espace des solutions. Cette comparaison est réalisée en considérant les deux codages sous leur forme matricielle en utilisant la notion de matrice de mise en correspondance.

Reprenons l'exemple de la matrice de mise en correspondance décrite dans la section 3.2 (cf. Figure 3.13). Considérons le chromosome entier de la figure 3.14 qui représente une solution de mise en correspondance entre les listes des points de contour  $G$  et  $D$  de la figure 3.13. Dans cette exemple, les listes  $T_{max}$  et  $T_{min}$  représentent les listes des points de contour  $D$  et  $G$ , respectivement. Comme nous l'avons indiqué plus haut, on peut transformer ce chromosome en un chromosome binaire, et donc le mettre sous une

forme matricielle où seuls les couples appariés, c'est à dire ceux ayant un état de mise en correspondance égal à 1, sont représentés. (cf. Figure 3.15).

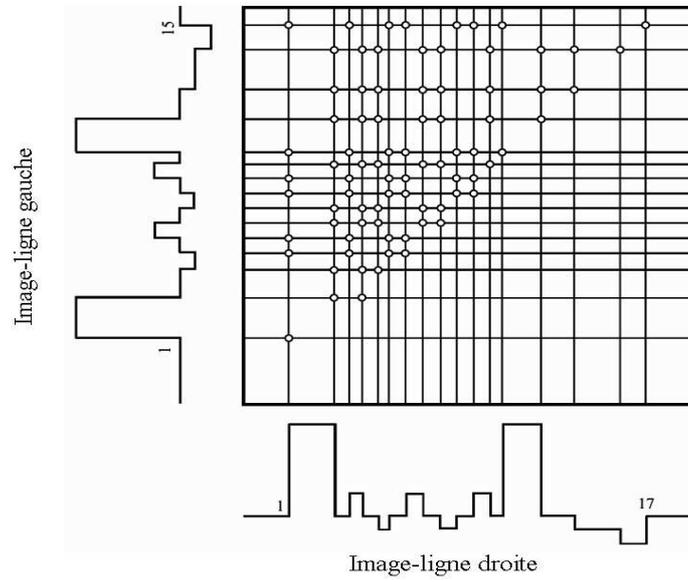


FIG. 3.13 – Exemple de matrice de mise en correspondance.

1	2	4	6	13	8	9	10	10	9	11	13	0	13	14	14	15
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

FIG. 3.14 – Exemple de chromosome entier.

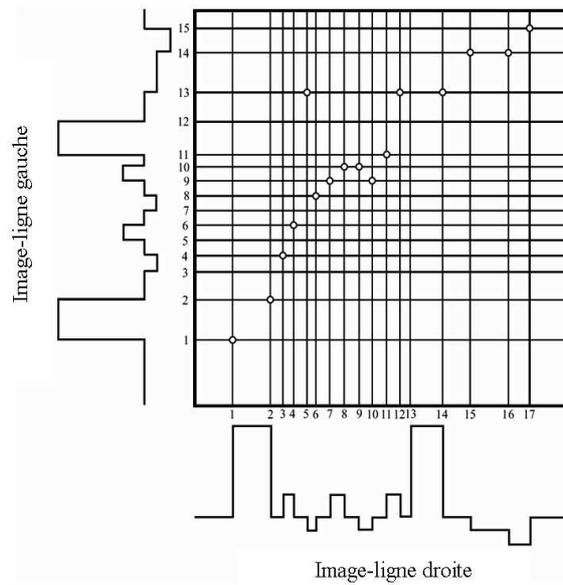


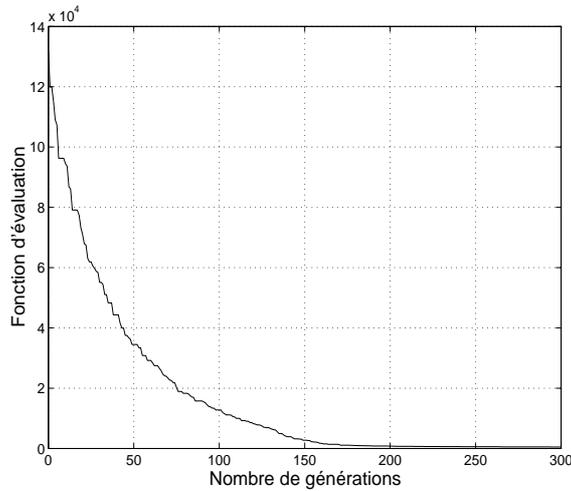
FIG. 3.15 – Forme matricielle du chromosome entier de la figure 3.14.

On peut constater que sur la forme matricielle du chromosome entier, les appariements multiples n'apparaissent que sur les lignes. c'est le cas pour tous les chromosomes entiers qui codent une solution de mise en correspondance entre les listes des points de contour gauche et droit de notre exemple (cf. Figure 3.15). Si on inverse les rôles de  $T_{min}$  et  $T_{max}$ , les appariements multiples n'apparaissent que sur les colonnes de la matrice de mise en correspondance. Les chromosomes entiers introduisent alors moins d'ambiguïtés vis-à-vis de la contrainte d'unicité que les chromosomes binaires, qui eux peuvent mettre en évidence des appariements multiples à la fois sur les lignes et sur les colonnes de la matrice de mise en correspondance.

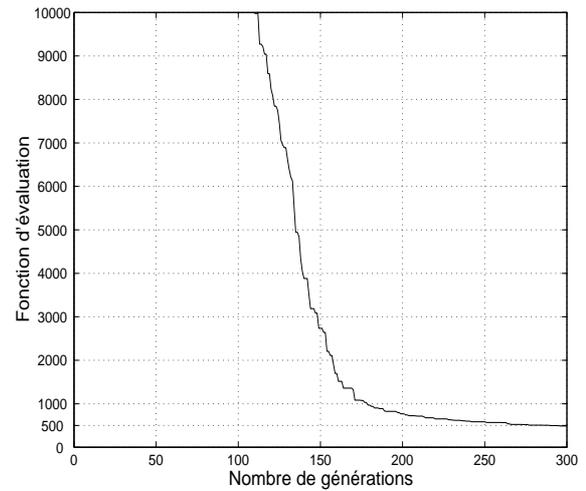
Partant de cette constatation, on peut supposer que le codage entier permet une meilleure exploration de l'espace des solutions. Pour vérifier cette hypothèse, nous avons utilisé deux algorithmes génétiques identiques à celui décrit dans la section 3.4. Le premier algorithme est un AGB, c'est à dire un algorithme génétique manipulant des chromosomes binaires qui sont évalués sous leur forme matricielle, en utilisant la fonction d'évaluation établie dans la section 3.3 (cf. Equation 3.1). Le deuxième algorithme utilise des chromosomes entiers et suit le même schéma que celui de l'AGB. Cet algorithme, baptisé algorithme génétique hybride (AGH), effectue toutes les opérations génétiques sur les

chromosomes entiers qui sont évalués eux aussi en considérant leur forme matricielle, avec la même fonction d'évaluation que celle de l'AGB (cf. Equation 3.1).

La figure 3.16 (resp. 3.17) représente l'évolution de la meilleure valeur d'adaptation durant l'exécution de l'AGB (resp. AGH). Ces résultats correspondent à l'application de ces deux algorithmes à un même couple d'images-ligne stéréoscopiques. Les deux algorithmes utilisent les mêmes valeurs pour tous les paramètres génétiques. En fixant la taille de la population  $Taille_{pop}$  à 100 chromosomes, nous constatons qu'au bout de 300 générations, l'AGH fournit un résultat nettement meilleur que celui obtenu par l'AGB. En effet, l'AGB aboutit à un "meilleur" chromosome avec une valeur d'adaptation d'environ 488 alors que le "meilleur" chromosome construit par l'AGH a une valeur d'adaptation d'environ  $-93$ . Nous rappelons que la recherche d'une bonne solution correspond à la minimisation de la fonction d'évaluation (cf. Equation 3.1). Par ailleurs, en examinant l'allure des deux évolutions, on constate qu'avec l'AGH on converge plus rapidement qu'avec l'AGB. En effet, on peut noter qu'au bout de 150 générations, la fonction d'évaluation atteint la valeur  $-15$  avec l'AGH, au lieu de 2735 avec l'AGB. En augmentant la taille de la population à 200 et le nombre de générations à 400, la performance de l'AGB s'améliore en mettant en évidence un chromosome avec une valeur d'adaptation égale à 209 (cf. Figure 3.18). Cependant, cette performance reste nettement inférieure à celle de l'AGH.



(a) Courbe en pleine échelle.



(b) Courbe avec échelle réduite.

FIG. 3.16 – Evolution de la fonction d'évaluation du meilleur chromosome par l'AGB, avec  $Taille_{pop} = 100$  et  $N_{gen} = 300$  pour le couple d'images-ligne de rang 135 de la séquence "Piéton".

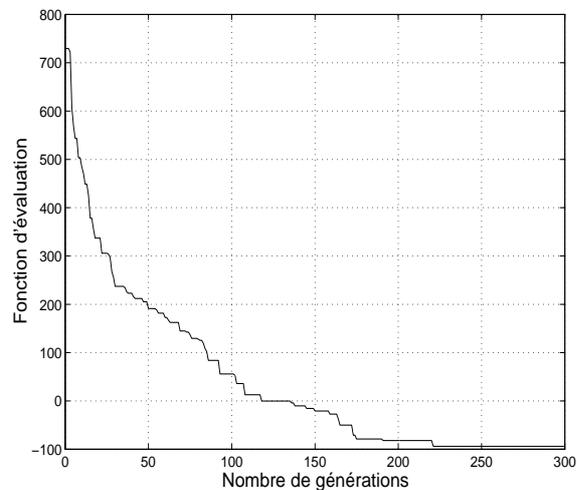
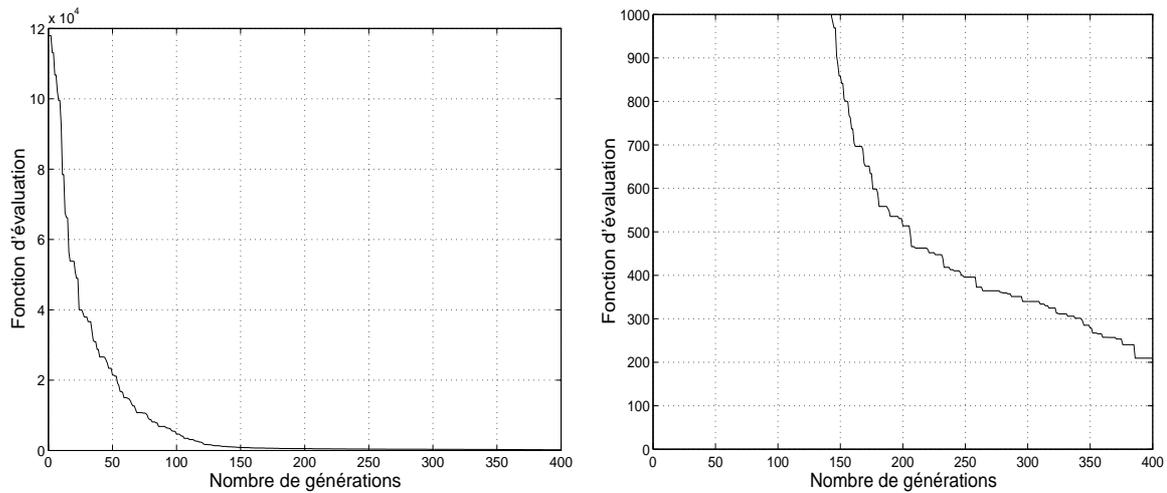


FIG. 3.17 – Evolution de la fonction d'évaluation du meilleur chromosome par l'AGH, avec  $Taille_{pop} = 100$  et  $N_{gen} = 300$  pour le couple d'images-ligne de rang 135 de la séquence "Piéton".



(a) Courbe en pleine échelle.

(b) Courbe avec échelle réduite.

FIG. 3.18 – Evolution de la fonction d'évaluation du meilleur chromosome par l'AGB, avec  $Taille_{pop} = 200$  et  $N_{gen} = 400$  pour le couple d'images-ligne de rang 135 de la séquence "Piéton".

Cette analyse permet de conclure que le codage entier permet une meilleure exploration de l'espace des solutions que le codage binaire. En plus, les chromosomes issus du codage entier sont de plus petites tailles. Dans l'exemple traité ci-dessus (cf. Figures 3.14 et 3.13), les chromosomes entiers sont composés de 17 gènes au lieu de 81 gènes pour les chromosomes binaires. Avoir des chromosomes de petites tailles permet de gagner du temps au niveau de certaines phases de l'algorithme génétique telles que l'évaluation et la mutation des chromosomes. Pour pouvoir réduire le temps des traitements, nous proposons, dans la section suivante, une nouvelle fonction permettant une évaluation directe des chromosomes entiers.

### 3.8 Evaluation des chromosomes entiers

Comme pour le codage binaire, la fonction d'évaluation des chromosomes entiers est construite à partir des contraintes d'unicité, d'ordre et de continuité de la disparité de telle sorte que sa valeur minimale corresponde à une solution où les appariements respectent

au mieux ces contraintes. Elle est exprimée sous la forme suivante (cf. Equation 3.3) :

$$\begin{aligned}
H_{entier} = & K_u \sum_{i \in T_{\max} \text{ et } i=1}^{N_{\max}-1} \sum_{j \in T_{\max} \text{ et } j=i+1}^{N_{\max}} U(C_i, C_j) \\
& + K_m \left( N_{\min} - \sum_{i \in T_{\max} \text{ et } i=1}^{N_{\max}} Z(C_i) \right)^2 \\
& + K_o \sum_{i \in T_{\max} \text{ et } i=1}^{N_{\max}-1} \sum_{j \in T_{\max} \text{ et } j=i+1}^{N_{\max}} O(C_i, C_j) \\
& - K_c \sum_{i \in T_{\max} \text{ et } i=1}^{N_{\max}-1} \sum_{j \in T_{\max} \text{ et } j=i+1 \text{ et } C_i \neq C_j}^{N_{\max}} S'(C_i, C_j) \quad (3.3)
\end{aligned}$$

où  $K_u$ ,  $K_m$ ,  $K_o$  et  $K_c$  sont des coefficients de pondération des différents termes de la fonction d'évaluation.

Le premier terme de la fonction d'évaluation correspond à la contrainte d'unicité. Le coefficient  $U(C_i, C_j)$  représente une pénalité quand deux points de contour d'une image ont le même correspondant. Il est exprimé de la manière suivante :

$$U(i, C_j) = \begin{cases} 1 & \text{si } i = C_j \text{ et } i \neq 0 \text{ et } C_j \neq 0 \\ 0 & \text{sinon} \end{cases}$$

Le deuxième terme est introduit pour renforcer les chromosomes qui mettent en évidence le maximum de couples appariés. Ce terme reste compatible avec la contrainte d'unicité puisque le nombre maximal de couples appariés est fixé à  $N_{\min} = \min(N_G, N_D)$ . Le coefficient  $Z(C_i)$  représente un indicateur de présence d'un couple apparié. Il est donné par :

$$Z(C_i) = \begin{cases} 1 & \text{si } C_i \neq 0 \\ 0 & \text{sinon} \end{cases}$$

Le troisième terme représente la contrainte d'ordre. Le coefficient  $O(C_i, C_j)$  indique, en terme de pénalité, si l'ordre entre les couples de points de contour  $(i, C_i)$  et  $(j, C_j)$  est

respecté ou non. Il est défini par :

$$O(C_i, C_j) = \begin{cases} 1 & \text{si } C_i > C_j \text{ et } C_i \neq 0 \text{ et } C_j \neq 0 \\ 0 & \text{sinon} \end{cases}$$

Le dernier terme de la fonction d'évaluation représente la contrainte de continuité de la disparité. Il est utilisé pour favoriser les chromosomes qui codent les couples ayant des disparités voisines. Le coefficient  $S'(C_i, C_j)$  représente une mesure de compatibilité entre les couples  $(i, C_i)$  et  $(j, C_j)$  vis-à-vis de la contrainte de continuité de la disparité. Il s'écrit sous la forme suivante :

$$S'(C_i, C_j) = S(X_{C_i, C_j}) = \frac{2}{1 + e^{\alpha(X_{C_i, C_j} - \theta)}} - 1 \quad (3.4)$$

où  $X_{C_i, C_j}$  est la valeur absolue de la différence des disparités des couples  $(i, C_i)$  et  $(j, C_j)$ .

La fonction  $S$  ainsi que les paramètres  $\theta$  et  $\alpha$  ont été décrits dans la section 3.3.

L'algorithme génétique utilisant le codage entier et la fonction d'évaluation décrite ci-dessus (cf. Equation 3.3) est baptisé algorithme génétique entier (AGE).

### 3.9 Réglage de la taille de la population et du nombre de générations

La taille de la population  $Taille_{pop}$  et le nombre de générations  $N_{gen}$  sont deux paramètres qui contrôlent le temps de convergence d'un algorithme génétique, c'est à dire le temps nécessaire pour aboutir à construire le "meilleur" chromosome. Leur choix est très important, car plus ils prennent des valeurs élevées plus le temps de convergence augmente. Nous avons constaté que le réglage de ces deux paramètres dépend de la complexité des images en terme de quantité de points de contour extraits : plus le nombre de points de contour augmente plus l'algorithme génétique a besoin d'une population de grande taille et d'un nombre important de générations pour aboutir à une solution acceptable. Même constatation quand la différence entre les nombres des points de contours extraits des deux images est importante.

Tenant compte de ces constatations, nous avons proposé une stratégie empirique pour régler la taille de la population et le nombre de générations. Elle est exprimée de la manière suivante :

$$Taille_{pop} = A \cdot (N_G + N_D) + B \cdot |N_G - N_D| \quad (3.5)$$

$$N_{gen} = C \cdot Taille_{pop} \quad (3.6)$$

où  $A$ ,  $B$  et  $C$  sont des constantes positives. Elles ont été fixées expérimentalement à 5, 2 et 1, respectivement.

### 3.10 Résultats de mise en correspondance obtenus par l'AGE

Les figures 3.19 et 3.20 représentent respectivement les images de reconstruction obtenues après application de l'AGE aux séquences stéréoscopiques "Piéton" et "Approche d'un véhicule".

Dans la section 3.7.2, nous avons montré quantitativement que le codage entier fournit des résultats meilleurs que le codage binaire. Qualitativement, on peut constater qu'il y a moins de faux appariements sur les images de reconstruction fournies par l'AGE (cf. Figures 3.19 et 3.4 – 3.20 et 3.5). Par contre, les résultats de l'AGE restent très comparables à ceux obtenus par la méthode neuronale (cf. Figures 3.19 et 3.6 – 3.20 et 3.7) qui est toujours la plus rapide (cf. Table 3.2).

A moins qu'on utilise une architecture parallèle à plusieurs processeurs, la cadence de traitement de l'AGE ne permet pas d'assurer la détection d'obstacles temps réel (cf. Table 3.2). Notre algorithme génétique pose également le problème de l'extraction des appariements à partir du "meilleur" chromosome. En effet, contrairement à la méthode neuronale, aucun traitement ne peut être effectué pour éliminer les éventuels appariements multiples. Ceci est dû au codage qui est de type "tout ou rien".

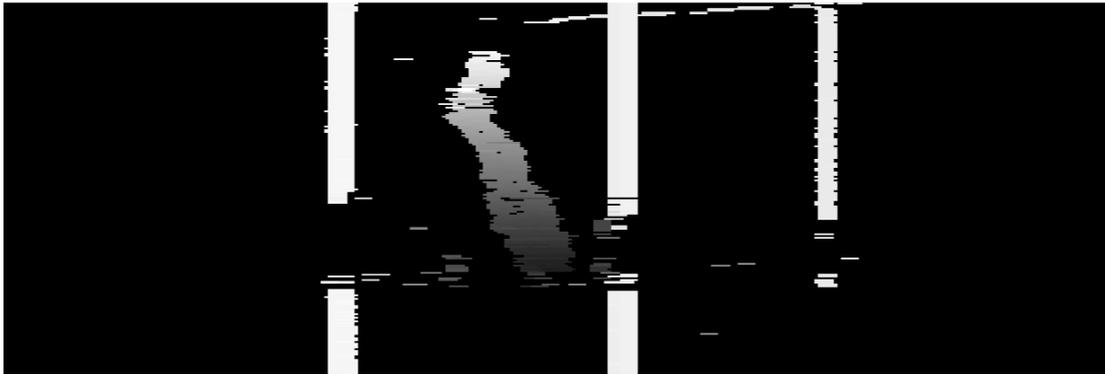


FIG. 3.19 – Séquence "Piéton" : Image de reconstruction obtenue par l'AGE.



FIG. 3.20 – Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par l'AGE.

TAB. 3.2 – Comparaison des cadences de traitement de l'AGE, de l'AGB et de la méthode neuronale

Méthode	Cadence de traitement (en nombre de couples d'images-ligne par seconde)	
	Séquence "Piéton"	Séquence "Approche d'un véhicule"
AGB	0.1	0.3
AGE	2.7	6.4
Méthode neuronale	90	256

## 3.11 Conclusion

Suite à nos travaux de recherche sur la mise en correspondance par les réseaux de neurones de Hopfield, nous avons voulu explorer l'apport des algorithmes génétiques qui sont des méthodes de recherche globales. Nous avons tout d'abord proposé un algorithme génétique basé sur un codage binaire classique. Nous avons rapidement constaté que ce codage n'était pas très approprié pour mieux explorer l'espace des solutions. Nous avons développé alors un nouveau codage, baptisé codage entier, fournissant des chromosomes plus compacts et permettant une meilleure exploration de l'espace des solutions.

Notre algorithme génétique n'apporte pas d'amélioration significative par rapport la méthode neuronale. En effet, les résultats de mise en correspondance des deux approches restent très comparables. En revanche, la méthode neuronale offre une cadence de traitement très supérieure à celle atteinte par la méthode génétique.

Cherchant toujours à répondre à la contrainte temps réel imposée par la détection d'obstacles à l'avant d'un véhicule, nous présentons dans le chapitre suivant deux approches permettant de réaliser une mise en correspondance rapide. Partant du fait que le problème de mise en correspondance est de nature combinatoire, la première approche consiste à le décomposer en plusieurs sous-problèmes de tailles réduites, en utilisant une stratégie de recherche hiérarchique. Ce travail a fait l'objet d'une collaboration avec J.-C. Burie, ancien chercheur de l'équipe "Image et Décision", actuellement, Maître de Conférences à l'Université de La Rochelle et membre du Laboratoire L3I. La deuxième approche considère le problème de mise en correspondance sous la forme d'un processus de vote où il s'agit d'affecter un score à chaque couple de points de contour. La stratégie de vote est régie par les contraintes du problème. Ce travail s'inscrit dans le cadre de la thèse de M. Hariti que j'ai co-encadrée avec A. Koukam au sein de l'équipe "Informatique" du Laboratoire Systèmes et Transports.

---

---

## CHAPITRE 4

---

Mise en correspondance rapide

## 4.1 Introduction

S'inscrivant dans le cadre de l'amélioration de la sécurité dans les systèmes de transports terrestres, le champ applicatif de nos travaux de recherche nous a toujours poussés à prendre en compte la contrainte temps réel imposée par certaines applications telles que la détection d'obstacles à l'avant d'un véhicule. Pour satisfaire cette contrainte, nous avons développé deux approches pour traiter le problème de mise en correspondance.

Partant du fait que ce problème est de nature combinatoire, la première approche consiste, dans un premier temps, à le décomposer en plusieurs sous-problèmes de combinaisons réduites à l'aide d'une stratégie de recherche hiérarchique. Ces sous-problèmes de mise en correspondance sont ensuite traités par les procédures basées sur les réseaux de neurones de Hopfield ou les algorithmes génétiques. Ce travail a été développé en collaboration avec J.-C. Burie, ancien chercheur de l'équipe "Image et Décision", actuellement Maître de Conférences à l'Université de La Rochelle. Il a donné lieu à deux publications dans des journaux et deux communications dans des conférences internationales.

La deuxième approche considère la tâche de mise en correspondance comme un processus de vote entre les couples de points de contour. Il s'agit d'attribuer un score à chaque couple en utilisant une stratégie de vote régie par les contraintes du problème. Ce travail, qui s'inscrit dans le cadre de la thèse de M. Hariti, a donné lieu à un chapitre de livre, quatre communications dans des conférences internationales et une communication dans une conférence nationale.

## 4.2 Stratégie de recherche hiérarchique

Le problème de mise en correspondance est de nature combinatoire (cf. Chapitre 1). Ainsi, plus la combinatoire est élevée, plus le coût de calcul des algorithmes de mise en correspondance augmente. Dans sa thèse, J.-C. Burie a présenté une stratégie de recherche hiérarchique pour mettre en correspondance les points de contour extraits des images ligne. Il a proposé cette stratégie de recherche dans le but de réduire les risques de propagation des erreurs pouvant survenir au cours de sa procédure de mise en correspondance

élémentaire, basée sur une analyse séquentielle des deux images images stéréoscopiques. Nous nous sommes rendus compte que cette stratégie permet de décomposer le problème de mise en correspondance en plusieurs sous-problèmes de combinatoires réduites. Nous avons alors décidé d'adapter cette stratégie à nos procédures de mise en correspondance, basées sur les réseaux de neurones de Hopfield et sur les algorithmes génétiques, afin d'améliorer leurs cadences de traitement.

Au lieu de mettre en correspondance simultanément tous les points de contour des deux images stéréoscopiques, l'idée de la recherche hiérarchique est de mettre en correspondance les points de contour en les considérant à plusieurs niveaux de recherche. Les couples de points de contour appariés à un niveau de recherche donné sont utilisés comme des couples de référence pour mettre en correspondance d'autres points de contour au niveau de recherche suivant.

La stratégie hiérarchique nécessite une procédure de sélection afin de déterminer les points de contour gauches et droits qui seront appariés à un niveau de recherche donné. Cette sélection doit être effectuée de telle sorte qu'un point de contour gauche (resp. droit), sélectionné à un niveau, ait son correspondant, s'il existe, parmi les points de contour droits (resp. gauches), sélectionnés au même niveau. Sans cette précaution, la recherche hiérarchique devient inutilisable à cause des erreurs qui peuvent être générées et, par conséquent, propagées tout au long du processus de mise en correspondance.

L'idée retenue pour définir le critère de sélection est que les contours délimitant des régions avec un fort contraste sont plus faciles à mettre en correspondance que les contours délimitant des régions présentant un faible contraste. Pour cela, nous utilisons l'amplitude du gradient des points de contour, obtenue lors de la phase d'extraction (cf. Chapitre 2). La stratégie de recherche hiérarchique consiste alors à mettre d'abord en correspondance les points de contour les plus significatifs, c'est à dire ceux ayant les plus fortes amplitudes, pour terminer avec les moins significatifs, c'est à dire ceux ayant de faibles amplitudes.

### 4.2.1 Principe de base

Pour des raisons de commodité d'appellation et de définition, nous allons nous servir de la notion de schéma de mise en correspondance à la place de la notion de matrice de mise en correspondance utilisée dans les chapitres 2 et 3. L'interprétation reste la même : un élément d'un schéma de mise en correspondance représente un couple de points de contour candidat à l'appariement, c'est à dire un couple de points de contour vérifiant les contraintes locales de position et de signe du gradient. Un schéma de mise en correspondance est défini à partir de deux couples de pixels des images-ligne gauche et droite. Il représente tous les couples candidats à l'appariement dont les points de contour sont situés entre les pixels des couples qui le définissent.

Considérons le schéma de mise en correspondance de la figure 4.1 qui représente tous les couples de points de contour candidats à l'appariement entre deux images-ligne stéréoscopiques. Ce schéma est défini à partir des couples de pixels  $(0, 0)$  et  $(N_P - 1, N_P - 1)$ , où  $0$  et  $N_P - 1$  sont les positions du premier et dernier pixel dans une image-ligne ayant  $N_P$  pixels. Ce schéma constitue le premier niveau de recherche. La première étape de la stratégie hiérarchique consiste à sélectionner dans les deux images, les points de contour les plus significatifs, c'est à dire ceux ayant les plus fortes amplitudes du gradient. La figure 4.2 illustre l'ensemble des couples candidats à l'appariement au premier niveau de recherche. A l'issue de la mise en correspondance des points de contour sélectionnés à ce niveau, nous obtenons une liste de couples appariés qui sert de base pour mettre en correspondance les points de contour les plus significatifs au deuxième niveau de recherche. En considérant deux à deux les couples successifs de cette liste, nous générons de nouveaux schémas de mise en correspondance qui sont traités au deuxième niveau (cf. Figure 4.3). Dans l'exemple de la figure 4.3, nous supposons que la mise en correspondance des points de contour sélectionnés au premier niveau de recherche fournit une liste de 4 couples appariés. Ainsi, 5 nouveaux schémas de mise en correspondances seront traités au deuxième niveau de recherche.

Au deuxième niveau de recherche, on applique le même procédé à chacun des nouveaux schémas de mise en correspondance. Le processus continue récursivement jusqu'à

ce qu'il n'y ait plus de points de contour à apparier.

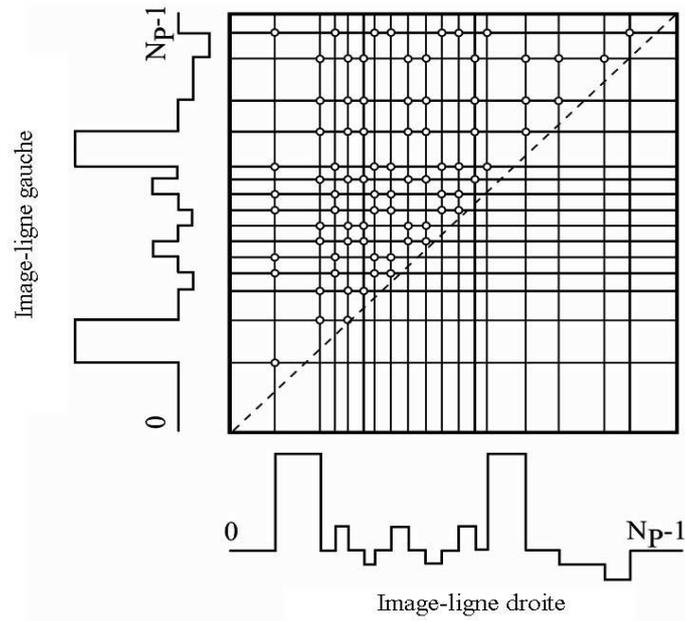


FIG. 4.1 – Schéma de mise en correspondance au premier niveau de recherche.

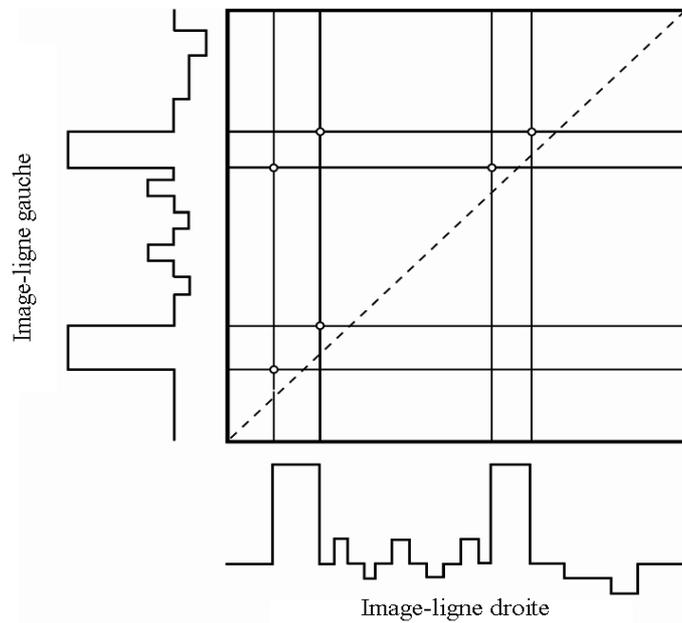


FIG. 4.2 – Couples candidats à l'appariement au premier niveau de recherche.

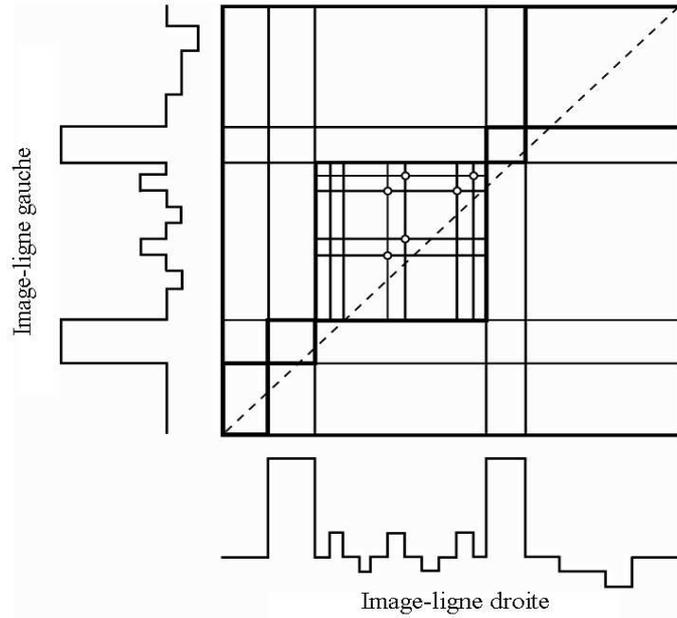


FIG. 4.3 – Schémas de mise en correspondance au deuxième niveau de recherche.

### 4.2.2 Définitions et formalisme

Comme nous l'avons indiqué dans la section 4.2.1, un schéma de mise en correspondance est défini à partir de deux couples de pixels. Chaque couple, appelé couple de référence, est constitué d'un pixel de l'image-ligne gauche et d'un pixel de l'image-ligne droite. Chacun de ces pixels peut correspondre ou non à un point de contour.

Pour pouvoir commencer le traitement, on définit le schéma de mise en correspondance au premier niveau de recherche grâce aux deux couples de référence  $(0, 0)$  et  $(N_P - 1, N_P - 1)$ . Ces deux couples de référence sont également utiles pour le traitement des nouveaux schémas de mise en correspondance situés aux extrémités du schéma initial. Par la suite, seuls les couples appariés à un certain niveau peuvent servir comme couples de référence pour générer des schémas de mise en correspondance au niveau supérieur.

Notons par  $CR_{n-1}^q(i) = (g_{n-1}^q(i), d_{n-1}^q(i))$  le  $i^{\text{ème}}$  couple de référence obtenu à partir du schéma de rang  $q$  au niveau  $n - 1$  où  $g_n^i(i)$  et  $d_n^q(i)$  sont respectivement deux pixels, représentant des points de contour ou non, des images-ligne gauche et droite. Les couples de référence  $CR_0^1(0) = (0, 0)$  et  $CR_0^1(1) = (N_p - 1, N_p - 1)$  qui servent à définir le schéma de mise en correspondance  $M_1^1(1) = [CR_0^1(0), CR_0^1(1)]$  au premier niveau sont supposés

être obtenus à partir d'un schéma fictif de rang 1 au niveau 0.

D'une manière générale, soit  $M_n^q(i) = [CR_{n-1}^q(i-1), CR_{n-1}^q(i)]$  le schéma de mise en correspondance de rang  $i$  au niveau  $n$  de recherche. Il est défini par les deux couples de référence successifs  $CR_{n-1}^q(i-1) = (g_{n-1}^q(i-1), d_{n-1}^q(i-1))$  et  $CR_{n-1}^q(i) = (g_{n-1}^q(i), d_{n-1}^q(i))$  obtenus à partir du schéma de rang  $q$  au niveau  $n-1$ . Le schéma  $M_n^q(i)$  ne prend en compte que les points de contour gauches (resp. droites) situés entre les pixels  $g_{n-1}^q(i-1)$  et  $g_{n-1}^q(i)$  (resp.  $d_{n-1}^q(i-1)$  et  $d_{n-1}^q(i)$ ) de l'image-ligne gauche (resp. droite). La liste de ces points de contour gauches (resp. droites) est notée par  $G_n^q(i)$  (resp.  $D_n^q(i)$ ).

Le traitement du schéma  $M_n^q(i)$  consiste, dans un premier temps, à sélectionner, à partir des listes  $G_n^q(i)$  et  $D_n^q(i)$ , les points de contour les plus significatifs. Notons par  $GS_n^q(i)$  et  $DS_n^q(i)$  les listes des points de contour sélectionnés à partir des listes  $G_n^q(i)$  et  $D_n^q(i)$ , respectivement. On applique ensuite la procédure de mise en correspondance aux listes  $GS_n^q(i)$  et  $DS_n^q(i)$ . On obtient alors une liste  $CR_n^i = \{CR_n^i(j), j = 0, \dots, J_n^i\}$  de  $J_n^i + 1$  couples de référence qui seront utilisés pour définir de nouveaux schémas de mise en correspondance au niveau  $n + 1$ . En plus des  $J_n^i - 1$  couples de points de contour appariés, cette liste contient également les couples de référence  $CR_{n-1}^q(i-1)$  et  $CR_{n-1}^q(i)$  qui définissent le schéma  $M_n^q(i)$ . Ces deux couples sont nécessaires à la génération des nouveaux schémas de mise en correspondance au niveau  $n + 1$ , notamment le premier et le dernier schémas. On a alors  $CR_n^i(0) = CR_{n-1}^q(i-1)$  et  $CR_n^i(J_n^i) = CR_{n-1}^q(i)$ .

A partir de la liste des couples de référence  $CR_n^i$ , on génère alors  $J_n^i$  nouveaux schémas au niveau  $n + 1$ . Notons par  $M_{n+1}^i = \{M_{n+1}^i(j), j = 1, \dots, J_n^i\}$  la liste de ces schémas où  $M_{n+1}^i(j) = [CR_n^i(j-1), CR_n^i(j)]$  est le schéma de mise en correspondance de rang  $j$  au niveau  $n + 1$ , obtenu à partir des  $(j-1)^{\text{ème}}$  et  $j^{\text{ème}}$  couples de référence issus du traitement du schéma de rang  $i$  au niveau  $n$ .

### 4.2.3 Sélection des points de contour

Comme nous l'avons expliqué plus haut, la stratégie hiérarchique consiste à mettre en correspondance les points de contour les plus significatifs, c'est à dire ceux qui présentent les plus fortes amplitudes du gradient, à chaque niveau de recherche.

Soit  $M_n^q(i) = [CR_{n-1}^q(i-1), CR_{n-1}^q(i)]$  le schéma de mise en correspondance de rang

$i$  au niveau  $n$  où  $CR_{n-1}^q(i-1) = (g_{n-1}^q(i-1), d_{n-1}^q(i-1))$  et  $CR_{n-1}^q(i) = (g_{n-1}^q(i), d_{n-1}^q(i))$ . Soit  $G_n^q(i)$  (resp.  $D_n^q(i)$ ) la liste des points de contour gauches (resp. droits) associés au schéma  $M_n^q(i)$ . Pour pouvoir procéder à la sélection des points de contour les plus significatifs à partir des listes  $G_n^q(i)$  et  $D_n^q(i)$ , nous supposons que celles-ci ne vérifient pas l'une des deux conditions d'arrêt de la recherche hiérarchique (cf. Section 4.2.5). Comme on s'intéresse au schéma  $M_n^q(i)$ , seuls les points de contour des listes  $G_n^q(i)$  et  $D_n^q(i)$  sont concernés par la sélection. Ces points de contour doivent alors vérifier la relation suivante :

$$x_{g_{n-1}^q}(i-1) < x < x_{g_{n-1}^q}(i), \text{ pour les points de contour de l'image-ligne gauche.} \quad (4.1)$$

$$x_{d_{n-1}^q}(i-1) < x < x_{d_{n-1}^q}(i), \text{ pour les points de contour de l'image-ligne droite.} \quad (4.2)$$

où  $x$  désigne l'abscisse des points de contour dans une image-ligne.

Soit  $d(x)$  l'amplitude de la dérivée associée au point de contour d'abscisse  $x$ . Le signal  $d(x)$  étant signé, nous prenons également en compte le signe de la dérivée. Appliquée à la liste  $G_n^q(i)$ , la procédure de sélection ne conserve que les points de contour dont la dérivée positive vérifie la relation :

$$\frac{Max_g}{2 \cdot n} \leq d(x) \leq Max_g \quad (4.3)$$

et les points de contour gauche présentant une dérivée négative vérifiant la relation :

$$Min_g \leq d(x) \leq \frac{Min_g}{2 \cdot n} \quad (4.4)$$

où  $Max_g$  (resp.  $Min_g$ ) est l'amplitude maximale (resp. minimale) de la dérivée de l'image-ligne gauche. La borne supérieure (resp. inférieure) de la relation 4.3 (resp. 4.4) étant fixe, cela permet à un point de contour n'ayant pas trouvé de correspondant au niveau  $n$  d'être considéré pour la mise en correspondance au niveau  $n+1$ .

Des relations similaires sont utilisées pour sélectionner les points les plus significatifs

de la liste  $D_n^q(i)$  :

$$\frac{Max_d}{2 \cdot n} \leq d(x) \leq Max_d \quad (4.5)$$

$$Min_d \leq d(x) \leq \frac{Min_d}{2 \cdot n} \quad (4.6)$$

où  $Max_d$  (resp.  $Min_d$ ) est l'amplitude maximale (resp. minimale) de la dérivée de l'image ligne droite.

A l'issue de la procédure de sélection, les points de contour de l'image gauche (resp. droite) sélectionnés constituent la liste  $GS_n^q(i)$  (resp.  $DS_n^q(i)$ ).

#### 4.2.4 Module d'appariement élémentaire

Soit  $M_n^q(i) = [CR_{n-1}^q(i-1), CR_{n-1}^q(i)]$  le schéma de mise en correspondance de rang  $i$  au niveau  $n$  où  $CR_{n-1}^q(i-1) = (g_{n-1}^q(i-1), d_{n-1}^q(i-1))$  et  $CR_{n-1}^q(i) = (g_{n-1}^q(i), d_{n-1}^q(i))$ . Soit  $G_n^q(i)$  (resp.  $D_n^q(i)$ ) la liste des points de contour gauches (resp. droits) associés au schéma  $M_n^q(i)$ . Selon le résultat du test d'arrêt effectué sur les listes  $G_n^q(i)$  et  $D_n^q(i)$  (cf. Section 4.2.5), le module d'appariement élémentaire peut être appliqué avec ou sans sélection.

Le module d'appariement avec sélection consiste, dans une première étape, à déterminer les listes des points de contour les plus significatifs  $GS_n^q(i)$  et  $DS_n^q(i)$  à partir des listes  $G_n^q(i)$  et  $D_n^q(i)$ , en utilisant les critères décrits dans la section 4.2.3. Les listes des points de contour sélectionnés  $GS_n^q(i)$  et  $DS_n^q(i)$  sont ensuite soumises à une procédure de mise en correspondance par réseau de neurones de Hopfield ou par algorithme génétique. Si on obtient une liste non vide de couples appariés alors on procède à la génération de nouveaux schémas de mise en correspondance qui seront traités de la même manière au niveau  $n+1$ . Dans le cas contraire, on continue la recherche avec le même schéma  $M_n^q(i)$ , mais au niveau  $n+1$ .

### 4.2.5 Conditions d'arrêt de la recherche hiérarchique

La procédure de recherche hiérarchique utilise deux conditions d'arrêt.

#### 4.2.5.1 Condition d'arrêt 1

Soit  $M_n^q(i) = [CR_{n-1}^q(i-1), CR_{n-1}^q(i)]$  le schéma de mise en correspondance de rang  $i$  au niveau  $n$  où  $CR_{n-1}^q(i-1) = (g_{n-1}^q(i-1), d_{n-1}^q(i-1))$  et  $CR_{n-1}^q(i) = (g_{n-1}^q(i), d_{n-1}^q(i))$ . Soit  $G_n^q(i)$  (resp.  $D_n^q(i)$ ) la liste des points de contour gauches (resp. droits) associés au schéma  $M_n^q(i)$ .

Si l'une des deux listes  $G_n^q(i)$  ou  $D_n^q(i)$  est vide, le module d'appariement élémentaire ne peut alors s'appliquer. Par conséquent, la recherche hiérarchique n'engendre plus de nouveaux schémas à partir du schéma  $M_n^q(i)$ .

#### 4.2.5.2 Condition d'arrêt 2

Il existe un niveau de recherche  $n_{max}$  à partir duquel nous sommes sûrs que tous les points de contour des deux images sont sélectionnés et traités par la procédure de recherche hiérarchique. La valeur  $n_{max}$  peut être fixée en fonction du critère de sélection (cf. Section 4.2.3) et du seuil  $S_b$  utilisé lors de l'extraction des points de contour (cf. Chapitre 2). On sait que la valeur absolue de l'amplitude du gradient des points de contour est supérieure à  $S_b$ . En utilisant l'équation 4.3, la procédure de sélection ne fournit plus de points de contour lorsque  $n_{max}$  vérifie la relation suivante :

$$\frac{Max_g}{2n_{max}} < S_b. \quad (4.7)$$

$n_{max}$  est donc le plus petit nombre entier vérifiant la relation :

$$\frac{Max_g}{2S_b} < n_{max}. \quad (4.8)$$

Des relations similaires peuvent être formulées à partir de  $Max_d$ ,  $Min_g$  et  $Min_d$  en prenant toutefois garde que pour les valeurs négatives  $Min_g$  et  $Min_d$  le seuil bas est  $-S_b$ .

La valeur maximale de l'amplitude de la dérivée est égale à 255 en valeur absolue.

Si on fixe le seuil  $S_b$  à 10, ceci nous conduit à une valeur de  $n_{max}$  égale à 13. Ainsi, si nous fixons  $n_{max}$  à 13, nous sommes sûrs de traiter tous les points de contour des deux images-ligne par la procédure de recherche hiérarchique. Le risque c'est que plus on monte dans les niveaux, plus la taille des schémas de mise en correspondance, c'est à dire le nombre de couples de points de contour candidats à l'appariement, diminue, et donc, plus l'apport des contraintes globales pour lever les ambiguïtés de mise en correspondance devient faible. Pour éviter ce problème, on propose d'ajuster  $n_{max}$  à des valeurs inférieures à 13. Dans ce cas, un schéma de mise en correspondance considéré au niveau de recherche  $n_{max}$  est traité par le module d'appariement élémentaire sans sélection. Si ce module fournit des couples appariés, on génère au niveau  $n_{max} + 1$  les nouveaux schémas à partir desquels aucun traitement n'est effectué. Dans le cas contraire, la recherche hiérarchique ne génère plus de nouveaux schémas.

#### 4.2.6 Implantation recursive de la stratégie hiérarchique

La stratégie de recherche hiérarchique est implantée selon un algorithme récursif afin de traiter tous les schémas de mise en correspondance, quel que soit leur nombre. Pour expliquer le fonctionnement de cet algorithme récursif, nous faisons appel à la notion d'arbre. Un arbre est un graphe constitué d'un ensemble de nœuds dont certaines paires sont connectées par des arcs orientés. Si un arc est orienté d'un nœud  $a$  vers un nœud  $b$ , on dit que le nœud  $a$  est le père du nœud  $b$  et que le nœud  $b$  est le fils du nœud  $a$ . Le nœud ne possédant pas de père constitue la racine de l'arbre. Un nœud n'ayant pas de fils constitue une feuille de l'arbre.

Dans notre cas, les nœuds de l'arbre sont les schémas de mise en correspondance. Le schéma de mise en correspondance du niveau de recherche 1 constitue la racine de l'arbre. A partir d'un nœud du niveau  $n$ , l'application du module d'appariement élémentaire (cf. Section 4.2.4) génère des nœuds fils au niveau  $n + 1$ . Un nœud de l'arbre peut être de type terminal, post-terminal ou normal. Un nœud est dit de type terminal, ou feuille, si le schéma qu'il représente vérifie la condition d'arrêt 1 (cf. Section 4.2.5.1). Un nœud est dit de type post-terminal si le schéma correspondant vérifie la condition d'arrêt 2 (cf. Section 4.2.5.2). Un nœud est dit de type normal s'il n'est ni de type terminal, ni de type

post-terminal.

#### 4.2.6.1 Construction de l'arbre de recherche

La procédure de recherche hiérarchique consiste à développer les nœuds de l'arbre en utilisant une exploration en profondeur d'abord, c'est à dire en traitant les nœuds de niveaux supérieurs avant de traiter les nœuds de même niveau. Lorsque tous les nœuds d'un même niveau sont traités, on remonte au niveau inférieur pour examiner les nœuds restants à ce niveau. Ainsi, lorsqu'un nœud du niveau  $n$  génère plusieurs nœuds fils au niveau  $n + 1$ , l'algorithme étend la recherche au niveau  $n + 2$ . Pour cela, une priorité est toujours donnée au nœud fils le plus à gauche.

Considérons un nœud au niveau de recherche  $n$ . Un test d'arrêt à deux conditions est effectué pour déterminer le type de ce nœud (cf. Section 4.2.5). Comme nous l'avons indiqué plus haut, le nœud peut être de type terminal, post-terminal ou normal.

Si le nœud est de type terminal, alors il cesse d'engendrer de nouveaux nœuds au niveau  $n + 1$ . Notons qu'un nœud de type terminal est un nœud feuille.

Dans le cas où le nœud est de type post-terminal, le traitement consiste à appliquer le module d'appariement élémentaire sans sélection (cf. Section 4.2.4). Celui-ci effectue la procédure de mise en correspondance sur tous les points de contour associés au schéma de mise en correspondance représenté par le nœud. Si ce module fournit des couples de points de contour appariés, on génère au niveau  $n + 1$  de nouveaux nœuds fils qui sont des nœuds feuille. Sinon, le nœud lui-même devient un nœud feuille.

Si le nœud est de type normal, on lance le module d'appariement élémentaire avec sélection. Celui-ci consiste à appliquer la procédure de mise en correspondance aux points de contour sélectionnés (cf. Section 4.2.4) à partir des listes des points de contour associées au schéma de mise en correspondance représenté par le nœud. Si ce module fournit des couples de points de contour appariés, on génère au niveau  $n + 1$  de nouveaux nœuds fils à partir desquels le processus est relancé. Dans le cas contraire, on re-traite le même nœud au niveau  $n + 1$ .

Le procédé de construction de l'arbre de recherche est illustré sur l'exemple de la figure 4.4. Cet exemple traite le schéma de mise en correspondance de la figure 4.1. On

commence par appliquer le module d'appariement élémentaire avec sélection au nœud racine  $M_1^1(1) = [CR_0^1(0), CR_0^1(1)]$ , avec  $CR_0^1(0) = (0, 0)$  et  $CR_0^1(1) = (N_p - 1, N_p - 1)$ . A l'issue du traitement de ce nœud, on obtient 4 couples de points de contour appariés. On constitue alors la liste des couples de référence  $CR_1^1$  :

$$CR_1^1 = \{CR_1^1(0) = CR_0^1(0), CR_1^1(1), CR_1^1(2), CR_1^1(3), CR_1^1(4), CR_1^1(5) = CR_0^1(1)\}.$$

A partir de cette liste, on crée 5 nouveaux nœuds fils au niveau de recherche 2 :  $M_2^1(1) = [CR_1^1(0), CR_1^1(1)]$ ,  $M_2^1(2) = [CR_1^1(1), CR_1^1(2)]$ ,  $M_2^1(3) = [CR_1^1(2), CR_1^1(3)]$ ,  $M_2^1(4) = [CR_1^1(3), CR_1^1(4)]$ ,  $M_2^1(5) = [CR_1^1(4), CR_1^1(5)]$ . Comme l'arbre est exploré en profondeur d'abord, on traite alors les nœuds  $M_2^1(1)$ ,  $M_2^1(2)$  puis  $M_2^1(3)$ . Les nœuds  $M_2^1(1)$  et  $M_2^1(2)$ , de type terminal, cessent d'engendrer de nouveaux nœuds. Le nœud  $M_2^1(3)$ , de type post-terminal, est traité par le module d'appariement élémentaire sans sélection. A partir des 4 couples de points de contour appariés, on constitue la liste des couples de référence  $CR_2^3$  :

$$CR_2^3 = \{CR_2^3(0) = CR_1^1(2), CR_2^3(1), CR_2^3(2), CR_2^3(3), CR_2^3(4), CR_2^3(5) = CR_1^1(3)\}.$$

Cinq nouveaux nœuds fils sont alors générés au niveau de recherche 3 :  $M_3^3(1)$ ,  $M_3^3(2)$ ,  $M_3^3(3)$ ,  $M_3^3(4)$ ,  $M_3^3(5)$  au niveau de recherche 3. Comme le nœud père est de type post-terminal, ces nœuds fils sont automatiquement des feuilles. Le processus continue de la même manière avec les nœuds  $M_2^1(4)$  puis  $M_2^1(5)$ . De type terminal, ces deux nœuds cessent d'engendrer de nouveaux nœuds.

#### 4.2.6.2 Extraction des couples appariés

Lorsque le processus cesse d'engendrer de nouveaux nœuds dans l'arbre, les couples de points de contour appariés se trouvent associés aux nœuds feuilles. Pour cela, il faut parcourir l'arbre de recherche construit pour atteindre les nœuds feuilles. Une autre manière, plus pratique, consiste à mémoriser les couples des points de contour appariés au fur et à mesure de la construction de l'arbre de recherche. Un appariement est établi dès qu'on

construit un nœud feuille dans l'arbre. Il suffit alors de conserver la paire correspondant au deuxième couple de référence qui définit le schéma de mise en correspondance associée au nœud feuille. Il faut toutefois ignorer le dernier nœud feuille de l'arbre, car le deuxième couple de référence  $(N_P - 1, N_P - 1)$  du schéma associé à ce nœud feuille ne constitue pas un couple de points de contour. Il a été utilisé avec le couple  $(0, 0)$  uniquement pour définir le schéma de mise en correspondance au niveau de recherche 1.

Dans le cas de l'exemple de la figure 4.4, après avoir traité le nœud racine  $M_1^1(1)$  au niveau de recherche 1, on construit au niveau 2 les feuilles  $M_2^1(1)$  puis  $M_2^1(2)$ . On mémorise alors successivement les couples  $CR_1^1(1)$  et  $CR_1^1(2)$ . On génère ensuite, à partir du nœud post-terminal  $M_2^1(3)$ , les feuilles  $M_3^3(1)$ ,  $M_3^3(2)$ ,  $M_3^3(3)$ ,  $M_3^3(4)$ ,  $M_3^3(5)$  au niveau 3, puis on récupère successivement les couples  $CR_2^3(1)$ ,  $CR_2^3(2)$ ,  $CR_2^3(3)$ ,  $CR_2^3(4)$ ,  $CR_2^3(5)$ . Finalement, on construit la feuille  $M_2^1(4)$  et on mémorise alors le couple  $CR_1^1(4)$ . Pour la dernière feuille  $M_2^1(5)$ , le couple  $CR_1^1(5)$  est ignoré puisqu'il ne correspond pas à un couple de points de contour.

## 4.2.7 Evaluation de la stratégie de recherche hiérarchique

La stratégie hiérarchique a été appliquée en considérant les procédures de mise en correspondance neuronale et génétique.

### 4.2.7.1 Séquence "Piéton"

La figure 4.5 représente l'image de reconstruction de la séquence "Piéton" obtenue par la méthode neuronale, appliquée d'une manière hiérarchique. Ces résultats de reconstruction sont comparables à ceux obtenus par application de la méthode neuronale sans la stratégie hiérarchique (cf. Figure 4.6). Sur l'image de la figure 4.5, on constate un peu plus de faux appariements qui nuisent à la reconstruction du piéton. Ces faux appariements apparaissent en particulier au milieu de la séquence où on peut noter la présence d'occultations dues au fait que le piéton cache la bande blanche de droite (resp. gauche) à la caméra gauche (resp. droite). Par rapport à l'image de reconstruction de la figure 4.6, ce résultat peut s'expliquer par un mauvais fonctionnement de la procédure de sélection utilisée dans

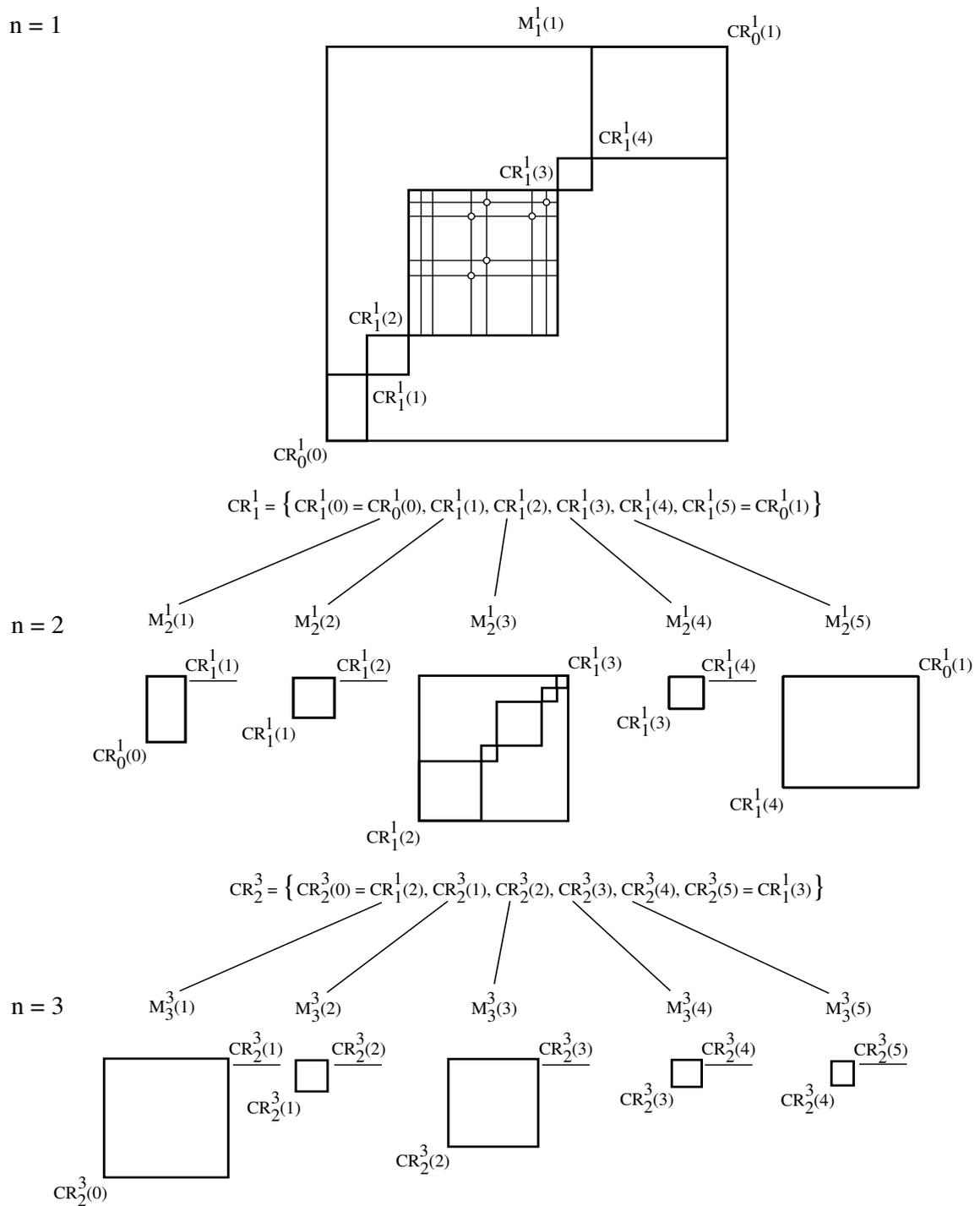


FIG. 4.4 – Arbre de recherche de la stratégie hiérarchique. Les couples soulignés représentent les couples de points de contour appariés.

la stratégie hiérarchique. En effet, à un niveau de recherche donné, cette procédure peut sélectionner un point de contour d'une image alors que son correspondant ne l'est pas dans l'autre image. Cette situation devient fréquente en particulier en présence d'occultations. On peut faire la même analyse quant aux résultats issus de la méthode génétique avec et sans la stratégie hiérarchique (cf. Figures 4.7 et 4.8). En revanche, la stratégie de recherche hiérarchique permet d'améliorer la cadence de traitement d'une manière très significative (cf. Table 4.1). Appliquée de manière hiérarchique, la méthode neuronale permet de traiter 260 paires d'images-ligne par seconde contre seulement 90 paires par seconde pour la même méthode sans la stratégie hiérarchique. Concernant la méthode génétique, la cadence de traitement avec la stratégie hiérarchique atteint 83 paires d'images-ligne par seconde au lieu de 2.7 paires par seconde sans la stratégie hiérarchique.

TAB. 4.1 – Séquence "Piéton" : Comparaison des cadences de traitement avec et sans recherche hiérarchique.

Méthode	Cadence de traitement (en nombre de couples d'images-ligne par seconde)	
	Avec la stratégie hiérarchique	Sans la stratégie hiérarchique
Neuronale	260	90
Génétique (AGE)	83	2.7



FIG. 4.5 – Séquence "Piéton" : Image de reconstruction obtenue par application hiérarchique de la méthode neuronale.

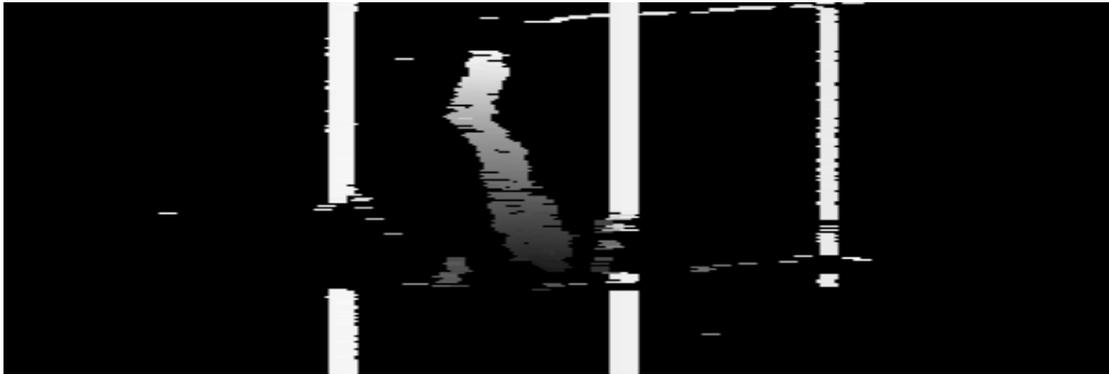


FIG. 4.6 – Séquence "Piéton" : Image de reconstruction obtenue par application de la méthode neuronale.

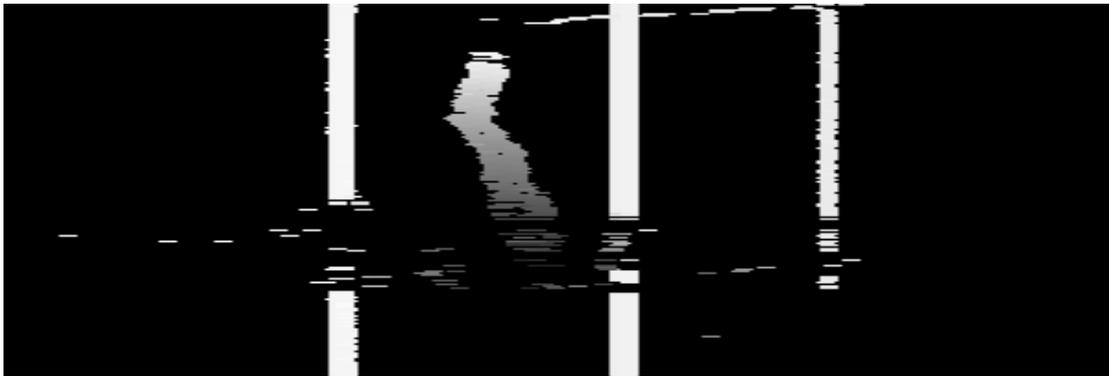


FIG. 4.7 – Séquence "Piéton" : Image de reconstruction obtenue par application hiérarchique de l'algorithme génétique entier.

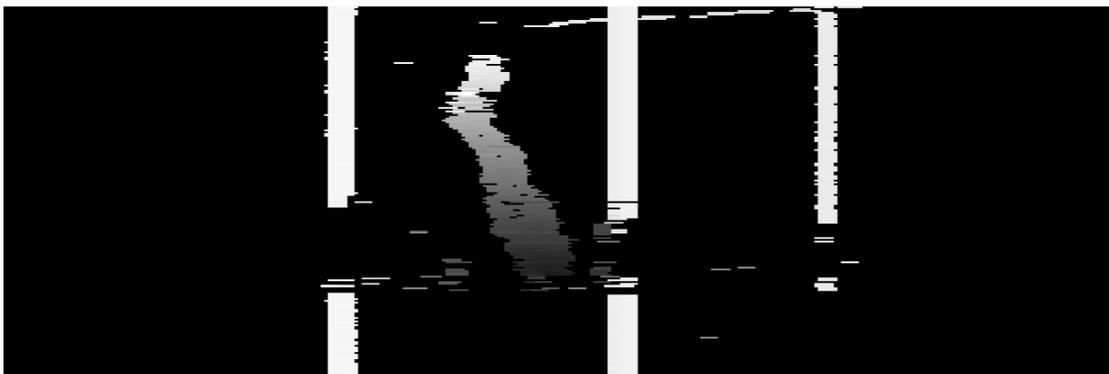


FIG. 4.8 – Séquence "Piéton" : Image de reconstruction obtenue par application de l'algorithme génétique entier.

### 4.2.7.2 Séquence "Approche d'un véhicule"

La figure 4.9 représente l'image de reconstruction issue du traitement de la séquence "Approche d'un véhicule" par la méthode neuronale hiérarchique. Ce résultat est comparable à celui obtenu avec la méthode neuronale sans la stratégie hiérarchique (cf. Figure 4.10). On remarque néanmoins que la stratégie de recherche hiérarchique provoque un peu plus d'erreurs d'appariement. On peut faire la même remarque sur les résultats obtenus à l'issue du traitement par la méthode génétique avec et sans la stratégie de recherche hiérarchique (cf. Figures 4.11 et 4.12). En revanche, la stratégie de recherche hiérarchique permet d'accélérer le temps de traitement des procédures de mise en correspondance neuronale et génétique (cf. Table 4.2). En effet, appliquée de manière hiérarchique, la méthode neuronale permet de traiter 328 paires d'images-ligne par seconde contre 256 paires par seconde pour la même méthode sans la stratégie hiérarchique. Concernant la méthode génétique, la cadence de traitement avec la stratégie hiérarchique atteint 68 paires d'images-ligne par seconde au lieu de 6.4 paires par seconde sans la stratégie hiérarchique.

TAB. 4.2 – Séquence "Approche d'un véhicule" : Comparaison des cadences de traitement avec et sans recherche hiérarchique.

Méthode	Cadence de traitement (en nombre de couples d'images-ligne par seconde)	
	Avec la stratégie hiérarchique	Sans la stratégie hiérarchique
Neuronale	328	256
Génétique (AGE)	68	6.4

### 4.2.8 Conclusion sur la recherche hiérarchique

Afin d'améliorer la cadence de traitement de nos procédures de mise en correspondance, nous avons présenté une stratégie de recherche hiérarchique permettant de décomposer le problème de mise en correspondance en plusieurs sous-problèmes de combinaisons réduites. Cette stratégie de recherche consiste à commencer par appairer les points de contour les plus significatifs, c'est à dire ceux qui présentent les plus fortes amplitudes du gradient, pour finir avec les moins significatifs, c'est à dire ceux qui présentent de

faibles amplitudes du gradient. A un niveau de recherche donné, on sélectionne les points de contour les plus significatifs, puis on les met en correspondance. Les couples de points de contour appariés servent de couples de référence pour mettre en correspondance les points de contours restant au niveau de recherche suivant, selon le même principe.

La stratégie de recherche hiérarchique permet de réduire significativement les cadences de traitement de nos procédures de mise en correspondance neuronale et génétique. En revanche, elle peut conduire à des erreurs d'appariement à cause de la sélection effectuée à chaque niveau de recherche, en particulier en présence d'occultations. En effet, un point de contour peut être sélectionné dans une image à un niveau donné alors que son correspondant ne l'est pas dans l'autre image.

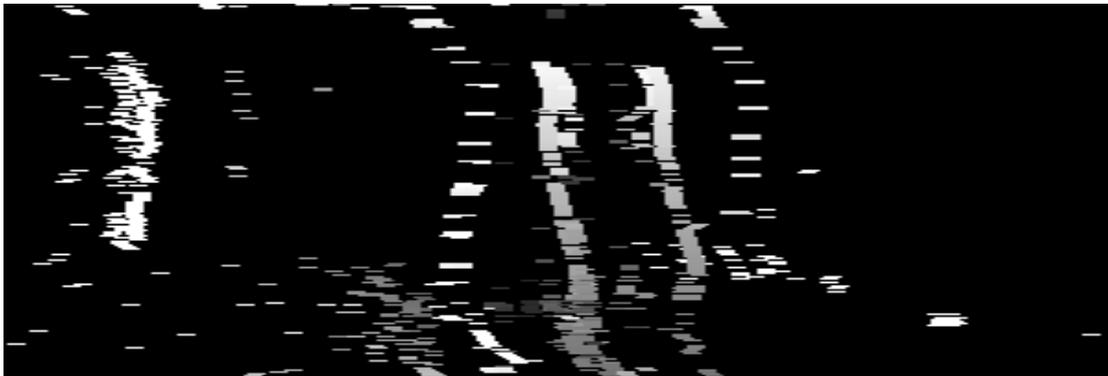


FIG. 4.9 – Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par application hiérarchique de la méthode neuronale.

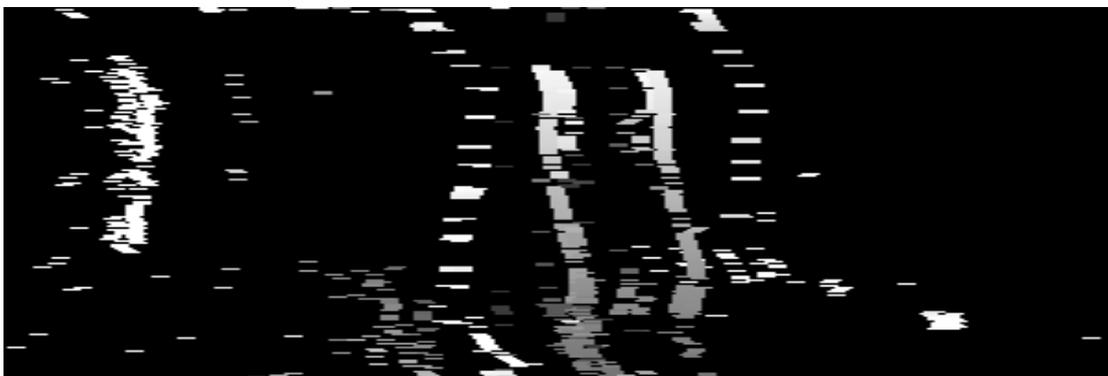


FIG. 4.10 – Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par application de la méthode neuronale.



FIG. 4.11 – Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par application hiérarchique de l'algorithme génétique entier.

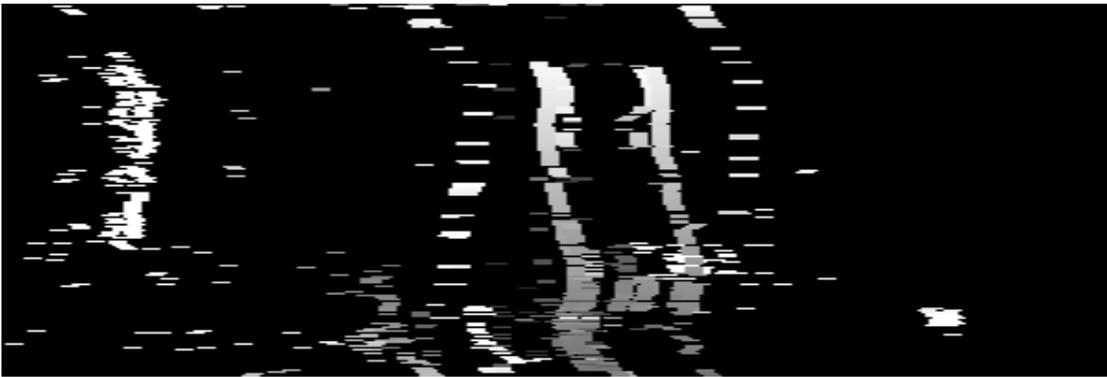


FIG. 4.12 – Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par application de l'algorithme génétique entier.

### 4.3 Stratégie de vote

Face aux limitations de la stratégie de recherche hiérarchique, nous avons développé une nouvelle approche qui considère le problème de mise en correspondance comme un processus de vote entre les couples de points de contour. Après élimination des couples qui ne vérifient pas les contraintes locales de position et de signe du gradient, une stratégie de vote, basée sur les contraintes globales, est mise en oeuvre pour affecter un score à chaque couple de points de contour.

Pour décrire la stratégie de vote, nous utilisons les termes couple "candidat" et

couple "électeur". Un couple candidat est un couple de points de contours qui respecte les contraintes locales. Un couple électeur d'un couple candidat est un couple de points de contour autorisé à contribuer au score du couple candidat.

Pour chaque couple candidat, la procédure de vote commence par déterminer ses couples électeurs en utilisant les contraintes d'unicité et d'ordre. Le score du couple est ensuite obtenu en additionnant les contributions des couples électeurs calculées grâce à la contrainte de continuité de la disparité. A l'issue de la procédure de vote, les appariements retenus sont ceux qui présentent les scores les plus élevés.

Pour expliquer la stratégie de vote, nous utilisons la notion de schéma de mise en correspondance introduite précédemment (cf. Figure 4.13). Nous rappelons qu'un élément  $M_{gd}$  de ce schéma représente une possibilité de mise en correspondance entre le point de contour  $g$  de l'image gauche et le point de contour  $d$  de l'image droite. Seuls les couples respectant les contraintes locales sont représentés. A chaque élément  $M_{gd}$ , nous associons une variable  $SM_{gd}$  représentant le score du couple  $(g, d)$ .

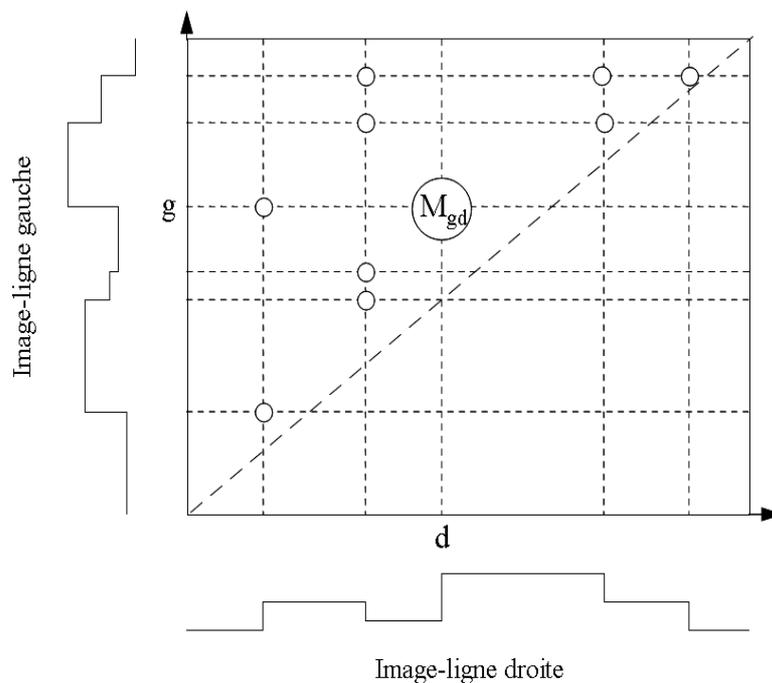


FIG. 4.13 – Schéma de mise en correspondance.

### 4.3.1 Etape 1 : Détermination des couples électeurs d'un couple candidat

Pour déterminer les couples électeurs d'un couple candidat, on utilise les contraintes globales d'unicité et d'ordre. Considérons un couple candidat  $(g, d)$ . Un couple  $(g', d')$  est considéré comme un couple électeur du couple  $(g, d)$  si les deux couples  $(g, d)$  et  $(g', d')$  vérifient les contraintes d'unicité et d'ordre. Cette première étape de la stratégie de vote est illustrée sur la figure 4.14. Pour satisfaire la contrainte d'unicité, on élimine tous les couples situés sur la ligne et sur la colonne associées à l'élément  $M_{gd}$ . La contrainte d'ordre permet ensuite d'écartier les couples situés sur les zones en haut à gauche et en bas à droite du schéma de mise en correspondance. Sur la figure 4.14, les couples électeurs sont situés à l'intérieur de la zone grise.

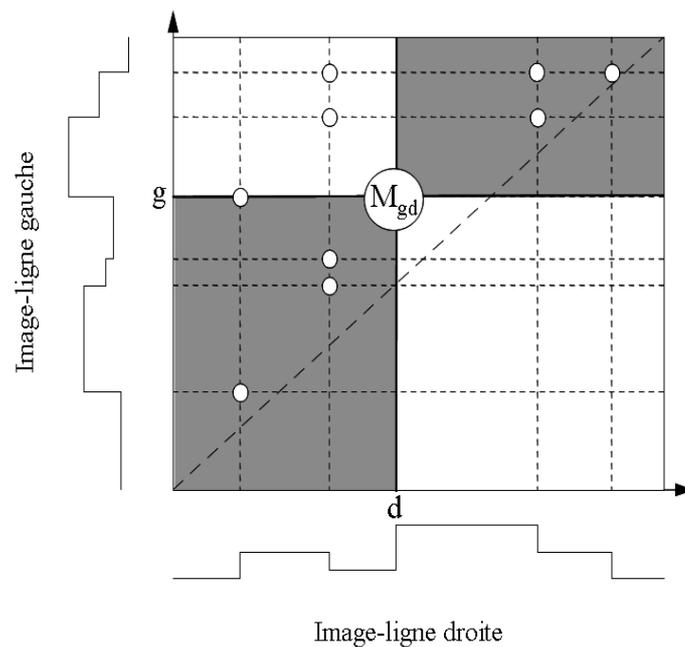


FIG. 4.14 – Détermination des couples électeurs d'un couple candidat.

### 4.3.2 Etape 2 : Calcul du score d'un couple candidat

Après avoir déterminé les couples électeurs d'un couple candidat, l'étape suivante consiste à calculer le score du couple candidat à partir du vote de ses couples électeurs en utilisant la contrainte de continuité de la disparité. Le vote de chaque couple électeur se traduit par sa contribution au score du couple candidat. Cette contribution quantifie la compatibilité entre les couples candidat et électeur vis-à-vis de la contrainte de continuité de la disparité. Elle est d'autant plus forte que les disparités des deux couples sont voisines.

Considérons un couple électeur  $(g', d')$  du couple candidat  $(g, d)$ . La contribution  $\Pi_{gdg'd'}$  du couple  $(g', d')$  est calculée de la manière suivante :

$$\Pi_{gdg'd'} = \Pi(X_{gdg'd'}) \quad (4.9)$$

où  $X_{gdg'd'}$  est la valeur absolue de la différence des disparités des couples  $(g, d)$  et  $(g', d')$  :

$$X_{gdg'd'} = |(x_g - x_d) - (x_{g'} - x_{d'})|$$

$x_g$ ,  $x_{g'}$ ,  $x_d$  et  $x_{d'}$  sont les positions des points de contour  $g$ ,  $g'$ ,  $d$  et  $d'$ , respectivement. La fonction  $\Pi$  est choisie de telle sorte qu'une compatibilité maximale, c'est à dire proche de 1, soit obtenue pour les couples qui correspondent à des valeurs de  $X_{gdg'd'}$  voisines de zéro, et qu'une compatibilité minimale, c'est à dire proche de zéro, soit obtenue pour les couples qui correspondent à des valeurs élevées de  $X_{gdg'd'}$ . Pour cela, nous avons utilisé une forme simple pour  $\Pi$  :

$$\Pi(X) = \frac{1}{1 + X}$$

A l'issue du vote de tous les couples électeurs, le score  $SM_{gd}$  du couple candidat  $(g, d)$  s'exprime alors de la façon suivante :

$$SM_{gd} = \sum_{(g', d') \in \Gamma_{gd}} \Pi_{gdg'd'} \quad (4.10)$$

où  $\Gamma_{gd}$  est l'ensemble des couples électeurs du couple candidat  $(g, d)$ , c'est à dire l'ensemble des couples qui vérifient les contraintes d'unicité et d'ordre avec le couple  $(g, d)$  :

$$\Gamma_{gd} = \{(g', d') \in \Omega / \text{les couples } (g, d) \text{ et } (g', d') \text{ respectent l'unicité et l'ordre}\}$$

où  $\Omega$  est l'ensemble des couples de points de contour qui vérifient les contraintes locales de position et de signe du gradient :

$$\Omega = \{(g', d') \in G \times D / (g', d') \text{ vérifie les contraintes locales}\}$$

$G$  et  $D$  sont les listes des points de contour gauches et droits, respectivement.

### 4.3.3 Extraction des appariements

Comme pour la méthode neuronale, l'extraction des couples appariés consiste à sélectionner les couples ayant les scores les plus importants sur les lignes et les colonnes du schéma de mise en correspondance. On cherche dans chaque ligne le couple ayant le score le plus élevé. Cette première recherche peut conduire à des situations où plusieurs couples sont choisis dans une même colonne. Pour éviter ces situations qui correspondent à des appariements multiples, on effectue la même procédure de recherche dans chaque colonne du schéma de mise en correspondance.

### 4.3.4 Evaluation de la stratégie de vote

La méthode de mise en correspondance par vote a été appliquée aux séquences d'images-ligne stéréoscopiques "Piéton" et "Approche d'un véhicule".

La figure 4.15 (resp. 4.18) représente l'image de reconstruction obtenue à l'issue du traitement de la séquence "Piéton" (resp. "Approche d'un véhicule") par la méthode de mise en correspondance par vote. Une comparaison avec les images de reconstruction obtenues par les méthodes neuronale et génétique montre que les résultats sont très comparables en terme de détection (cf. Figures 4.16, 4.17, 4.19 et 4.20). En revanche, la méthode de vote présente l'avantage d'être la plus rapide (cf. Table 4.3). En effet, pour la

séquence "Piéton", la cadence de traitement de la méthode de vote s'élève à 220 couples images-ligne par seconde contre 90 couples par seconde pour la méthode neuronale et 2.7 couples par seconde pour la méthode génétique. Pour la séquence "Approche d'un véhicule", la méthode de vote permet de traiter 364 couples par seconde au lieu de 256 couples par seconde pour la méthode neuronale et 6.4 couples par seconde pour la méthode génétique.

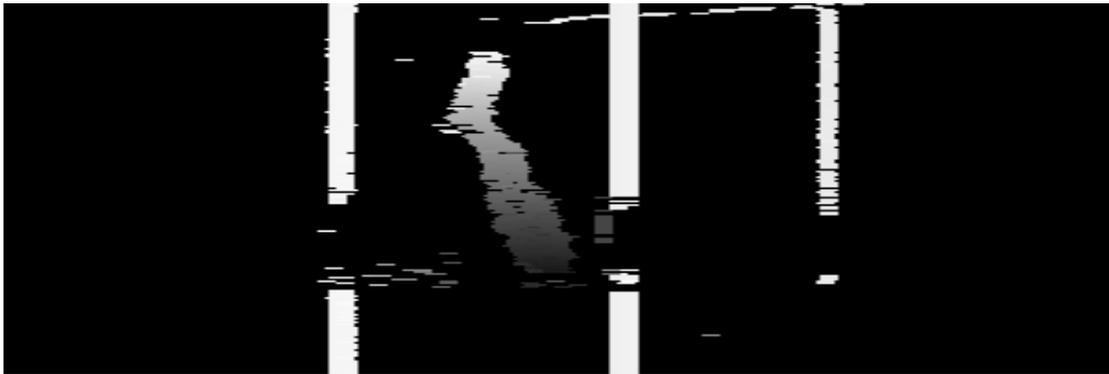


FIG. 4.15 – Séquence "Piéton" : Image de reconstruction obtenue par la méthode de vote.

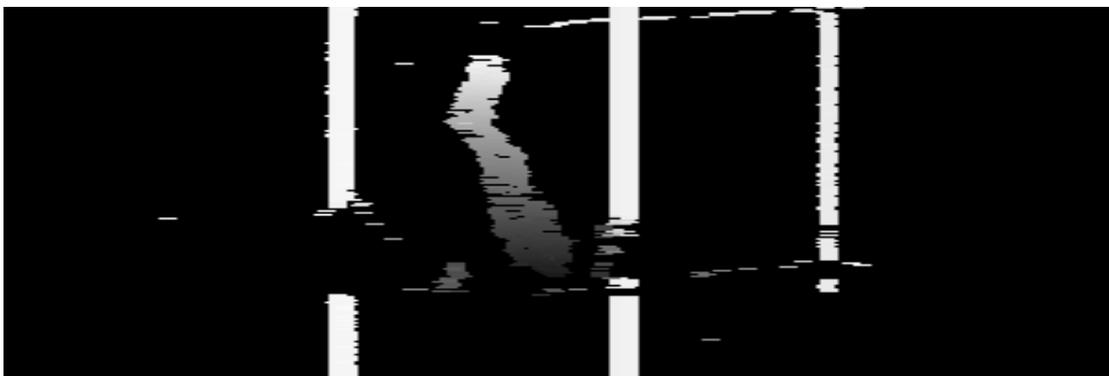


FIG. 4.16 – Séquence "Piéton" : Image de reconstruction obtenue par la méthode neuronale.

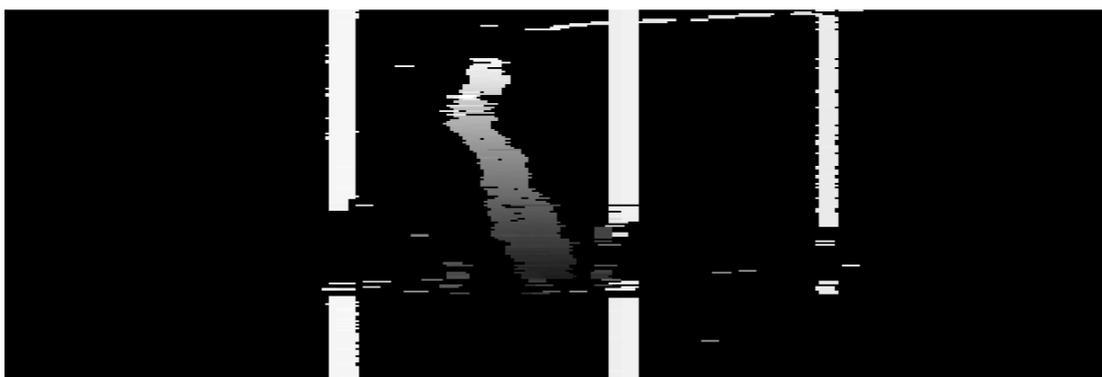


FIG. 4.17 – Séquence "Piéton" : Image de reconstruction obtenue par l'algorithme génétique entier.

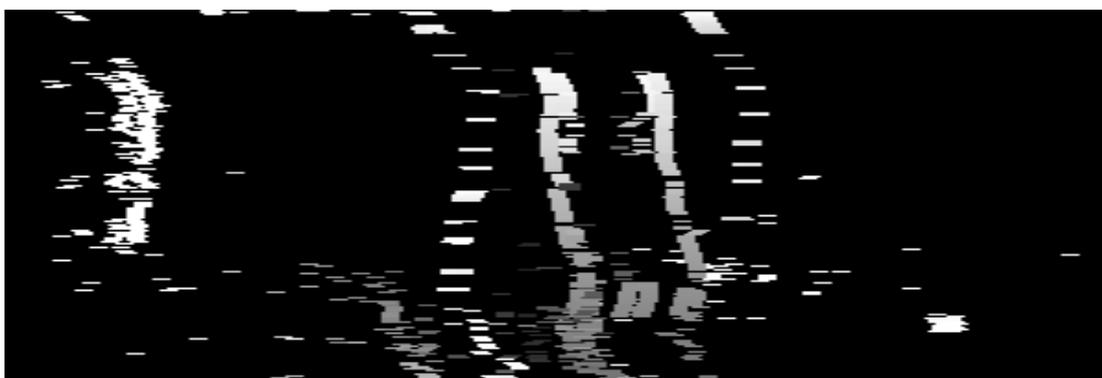


FIG. 4.18 – Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par la méthode de vote.

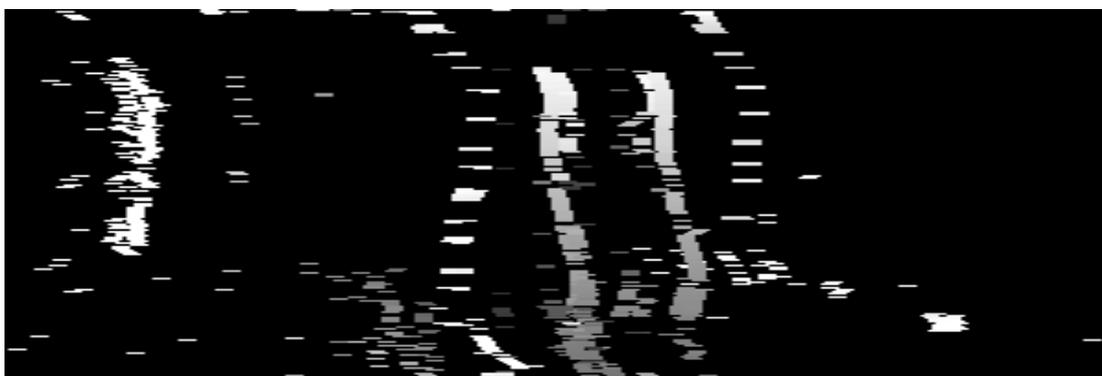


FIG. 4.19 – Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par la méthode neuronale.

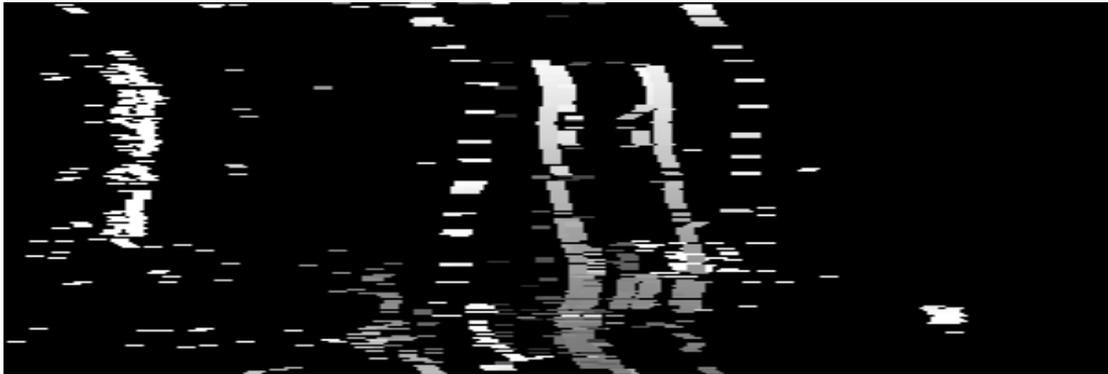


FIG. 4.20 – Séquence "Approche d'un véhicule" : Image de reconstruction obtenue par l'algorithme génétique entier.

TAB. 4.3 – Comparaison des cadences de traitement des méthodes de vote, neuronale et génétique.

Méthode	Cadence de traitement (en nombre de couples d'images-ligne par seconde)	
	Séquence "Piéton"	Séquence "Approche d'un véhicule"
Vote	220	364
Neuronale	90	256
Génétique	2.7	6.4

### 4.3.5 Conclusion sur la méthode de vote

Afin d'éviter les limitations de la stratégie de recherche hiérarchique, nous avons développé une nouvelle méthode de mise en correspondance permettant d'atteindre une cadence de traitement bien meilleure que celles des méthodes neuronale et génétique. Basée sur une stratégie de vote, cette méthode consiste à affecter un score à chaque couple de points de contour vérifiant les contraintes locales de position et de signe du gradient. La procédure de vote est régie par les contraintes globales d'unicité, d'ordre et de continuité de la disparité. Pour chaque couple candidat, on commence par déterminer ses couples électeurs par application des contraintes d'unicité et d'ordre. Le score du couple candidat est ensuite calculé en considérant la contribution de chaque couple électeur.

Cette contribution quantifie la compatibilité des couples candidat et électeur vis-à-vis de la contrainte de continuité de la disparité. A l'issue du processus de vote, on retient les couples ayant les scores les plus importants.

En plus de ses performances tant au niveau de la qualité des résultats qu'au niveau de la cadence de traitement, la méthode de vote ne fait intervenir aucun paramètre. Cette caractéristique lui confère l'avantage d'éviter les difficultés liées à l'ajustement de paramètres qui interviennent dans la plupart des méthodes.

## 4.4 Conclusion

Dans ce chapitre, nous avons présenté deux approches pour répondre à la contrainte temps réel imposée par le contexte applicatif de nos travaux de recherche.

Le problème de mise en correspondance étant de nature combinatoire, la première approche consiste à le décomposer en plusieurs sous-problèmes de combinatoires réduites. Pour cela, nous avons utilisé une stratégie de recherche hiérarchique permettant de mettre en correspondance les points de contour à plusieurs niveaux, des plus significatifs vers les moins significatifs vis-à-vis de l'amplitude de leur gradient. A un niveau donné, on commence par sélectionner les points de contour les plus significatifs. A l'issue de l'appariement de ces points de contour, les couples appariés sont utilisés comme couples de référence pour mettre en correspondance les points de contour restants au niveau supérieur, selon le même principe. Cette stratégie permet d'améliorer considérablement la cadence de traitement de nos procédures de mise en correspondance basées sur les réseaux de neurones de Hopfield et les algorithmes génétiques. En revanche, elle peut conduire à des erreurs d'appariement à cause de la sélection des points de contour significatifs effectuée à chaque niveau de recherche. En effet, celle-ci peut sélectionner un point de contour dans une image et ne pas sélectionner son correspondant dans l'autre image.

Afin d'éviter ce problème lié à la stratégie hiérarchique, la deuxième approche aborde le problème de mise en correspondance sous la forme d'un processus de vote qui affecte un score à chacun des couples de points de contour respectant les contraintes locales de position et de signe du gradient. Pour chaque couple candidat, la stratégie de vote com-

---

mence par déterminer ses couples électeurs en appliquant les contraintes globales d'unicité et d'ordre. Le vote d'un couple électeur se traduit par sa contribution au score du couple candidat. Cette contribution est obtenue en évaluant la compatibilité des couples candidat et électeur vis-à-vis de la contrainte de continuité de la disparité. Le score du couple candidat est ensuite calculé en additionnant les contributions de ses couples électeurs. Les couples appariés correspondent à ceux ayant les plus forts scores. Cette méthode permet d'obtenir des résultats d'appariement très comparables à ceux obtenus par les méthodes neuronales et génétiques, tout en présentant l'avantage d'être la plus rapide avec une cadence de traitement mieux adaptée à la contrainte temps réel. En plus, la méthode de vote ne nécessite aucun ajustement de paramètre. Il s'agit d'un avantage déterminant par rapport à la plupart des autres méthodes pour lesquelles la qualité des résultats dépend en général fortement de l'ajustement correct d'un ou même de plusieurs paramètres.

---

## Conclusion et perspectives

---

## Bilan

Dans ce mémoire, j'ai présenté mes activités de recherche qui se sont déroulées au laboratoire LAGIS, depuis 1993, et au laboratoire SeT, depuis 2001. Je me suis intéressé à la perception de l'environnement par stéréovision et à son application à la sécurité dans les systèmes de transports terrestres. Nous avons abordé (le terme "nous" s'impose car il souligne le caractère collectif de ces activités et des idées qui s'en sont dégagées) le problème de mise en correspondance d'images stéréoscopiques avec une attention particulière à la contrainte temps réel imposée par des applications telles que la détection d'obstacles à l'avant d'un véhicule. Depuis leur commencement, nos travaux de recherche se sont enchaînés de façon logique ! En effet, l'approche neuronale nous a conduits à remettre en cause la nature locale du processus de minimisation induit par un réseau de neurones de Hopfield. Nous avons alors proposé d'utiliser les algorithmes génétiques qui sont des méthodes de recherche globale. Le codage binaire classique ayant montré ses limites, nous avons proposé un nouveau codage, le codage entier, permettant une meilleure exploration de l'espace des solutions. En plus de son apport peu significatif par rapport à l'approche neuronale, l'approche génétique, fonctionnant avec une faible cadence de traitement, n'est pas compatible avec la contrainte temps réel. Cette limitation des algorithmes génétiques nous a poussés à réfléchir à une stratégie permettant d'améliorer le temps de calcul de nos procédures de mise en correspondance. Le problème de mise en correspondance étant de nature combinatoire, nous avons ainsi proposé de le décomposer en exploitant la stratégie de recherche hiérarchique qui consiste à apparier les points de contours des images à plusieurs niveaux. Cette technique permet d'améliorer significativement les cadences de traitement de nos procédures, basées sur les réseaux de neurones de Hopfield et sur les algorithmes génétiques. En revanche, elle se trouve confronté au risque de l'incompatibilité de la sélection des points de contour gauches et droits à chaque niveau de recherche. En effet, à un niveau de recherche donné, un point de contour peut être sélectionné dans une image alors que son correspondant, s'il existe, ne l'est pas dans l'autre image. Face à ce problème, et en cherchant toujours à mieux répondre à la contrainte temps réel, nous avons proposé une autre approche de mise en correspondance basée sur la technique

---

de vote. Cette approche permet d'atteindre des cadences de traitement bien meilleures avec une qualité des résultats très comparable à celles issues des approches neuronale et génétique.

## Perspectives

### *Continuation des travaux précédents*

Dans le cadre de mes collaborations avec le Laboratoire LAGIS et l'INRETS, ces travaux sont actuellement exploités pour le développement d'un système stéréoscopique pour le comptage de personnes entrant et sortant d'un bus. Il s'agit de localiser et suivre les têtes de personnes en combinant la stéréovision éparsée et dense. Ce travail, qui s'inscrit dans le cadre d'un contrat CIFRE avec la RATP, fait l'objet de la thèse de T. Yahiaoui que je co-encadre avec F. Cabestaing du laboratoire LAGIS et L. Khoudour de l'INRETS.

Depuis mon arrivée au laboratoire SeT, je mène le projet de recherche "Véhicule Intelligent pour la Ville du Futur". Ce projet me conduit naturellement à exploiter mes travaux précédents sur la perception de l'environnement. Une des problématiques qui me tient à coeur est la combinaison des perceptions des environnements avant, arrière et latéraux. Ce sujet sera traité dans le cadre d'une thèse qui débutera en octobre 2005. Cette thèse, en co-tutelle avec l'Université des Sciences et Technologies de Cracovie, a bénéficié d'une bourse de l'ambassade de France en Pologne.

J'ai commencé à travailler également sur la détection et le suivi d'objets en mouvement par utilisation des modèles de contours actifs. Nous considérons le cas où les objets entrent et sortent par la périphérie des images. La méthode la plus utilisée dans ce cas est d'installer un contour actif à la périphérie des images et détecter les objets en leur associant chacun un contour actif. Le suivi des objets est en suite mis en oeuvre en se basant sur l'estimation du mouvement des objets. En effet, à partir de sa position dans l'image courante, le contour actif de chaque objet est re-positionné par estimation pour détecter l'objet dans l'image suivante. Le problème de cette méthode est le risque de perdre les objets en présence d'occultations ou en cas de changement brusque de mouvement. Pour éviter ce problème, nous proposons d'effectuer la phase de détection pour chaque image

---

de la séquence, c'est à dire détecter les objets dans chaque image, en utilisant un contour actif de type "ballon" évoluant du bas vers le haut de l'image. A l'issue du traitement de chaque image de la séquence, un contour actif est associé à chaque objet. Le suivi d'objets est effectué ensuite par mise en correspondance des contours actifs. Cette méthode fonctionne sans tenir compte du type de mouvement des objets. Grâce à la détection effectuée pour chaque image, un objet perdu lors d'une occultation peut être retrouvé dès qu'il réapparaît sur les images. Les premiers résultats de ces travaux ont été présentés dans le cadre d'un DEA en 2004.

### *Navigation autonome*

Le projet "Véhicule Intelligent pour la Ville du Futur" m'ouvre l'opportunité de travailler sur de nouvelles thématiques que je souhaite aborder à court et à moyen terme. Ces thématiques, liées à la navigation autonome d'un véhicule dans un environnement urbain, portent sur deux problèmes très connus : la localisation de véhicules autonomes et la planification de leur mouvement. Pour le problème de la localisation, je propose une approche basée sur la vision artificielle et la réalité virtuelle. L'idée est de combiner trois sources d'informations : un système GPS (Global Positioning System), un système de vision et un système de simulation utilisant la réalité virtuelle. Le système de vision consiste à comparer les images issues de caméras embarquées avec des images géo-référencées générées automatiquement en temps réel à partir d'un système d'information géographique (SIG). Connaissant le trajet à parcourir, le système de simulation est destiné à produire dans un monde virtuel la navigation autonome du véhicule en supposant que le trajet est dégagé. L'objectif de ce système est de prédire à chaque instant la position du véhicule. Pour pouvoir utiliser les informations de localisation issues du système de simulation, il est donc nécessaire de synchroniser la navigation dans le monde réel et celle dans le monde virtuel. Pour cela, il faut gérer la simulation en fonction de l'environnement perçu dans le monde réel. Ainsi, l'arrêt du véhicule dans le monde réel, à cause d'un obstacle détecté par exemple, doit entraîner l'arrêt de la navigation virtuelle jusqu'à ce qu'il n'y ait plus d'obstacle.

Une deuxième approche un peu plus ambitieuse consiste à guider le véhicule dans le

monde réel en utilisant la navigation dans le monde virtuel. Le véhicule est amené à être contrôlé par la navigation simulée dans un monde virtuel. N'ayant aucune information de l'environnement réel, celle-ci doit être contrôlée par le système de perception associé au véhicule. Les deux mondes réel et virtuel doivent donc coopérer pour se synchroniser.

Nous prévoyons de commencer à aborder ces problèmes dès septembre 2005 dans le cadre d'une thèse pour laquelle nous avons demandé une bourse auprès du Conseil Régional de Franche Comté. Cette demande a été classée 2<sup>ème</sup> par le Conseil Scientifique de l'UTBM.

### *Modélisation multi-agents en analyse d'images*

Depuis mon arrivée au sein de l'équipe "Informatique" du laboratoire SeT, j'ai créé une nouvelle synergie en associant les systèmes multi-agents, un des domaines de compétence de cette équipe, et l'analyse d'images. Il s'agit de développer des modèles d'agents intelligents pour résoudre des problèmes en vision artificielle tels que la détection de contour, la segmentation, la modélisation d'objets par les modèles des contours actifs et la mise en correspondance des images. Récemment, nous avons développé une approche multi-agents pour la segmentation d'images. Grâce à leurs comportements individuel et collectif, des agents explorent un environnement (une image dans notre cas) pour chercher les contours des objets présents dans l'image. Les résultats, comparés à d'autres algorithmes de segmentation, se sont révélés très intéressants tant au niveau de la qualité de la segmentation qu'au niveau du temps de calcul.

---

## Annexe 1

### Activités d'enseignement à l'UTBM

## Bilan par matière

Les volumes horaires indiquent une moyenne sur l'ensemble des années.

### **Traitement et analyse d'images**

Cette unité de valeur a pour objectif de former les futurs ingénieurs aux outils théoriques et pratiques nécessaires à l'exploitation des images numériques, depuis l'acquisition jusqu'à la segmentation. Les principaux points traités sont :

- L'acquisition d'une image
- La topologie dans une image
- Les transformations d'images
- La réduction du bruit par filtrage
- L'extraction des contours
- La Segmentation

Filière ICI (département GI), en 2002 : 32 h cours, 30 h TD, 48 h TP

Filière I2RV (département GI), depuis 2003 : 26 h cours, 40 h TD, 44 h TP

### **Traitement avancé de l'information**

Cette unité de valeur traitait les nouvelles techniques de l'intelligence artificielle destinées à l'ingénierie de la décision. Ont été abordés :

- La logique floue
- Les réseaux de neurones
- Les algorithmes génétiques

Filière I2RV (département GI), en 2003 : 24h cours, 44 h TD, 48 h TP

### **Reconnaissance des formes**

Cette unité de valeur remplace la précédente (Traitement avancé de l'information). L'objectif est de donner aux élèves-ingénieurs les concepts nécessaires au développement de

---

systèmes intelligents capables de reproduire certaines fonctions de reconnaissance du cerveau à partir de données issues de la perception visuelle. Le programme comprend :

- La représentation d'une forme
- La reconnaissance basée sur la logique floue
- La reconnaissance basée sur les réseaux de neurones

Filière I2RV (département GI), depuis 2004 : 20 h cours, 18 h TD

### **Vision artificielle**

Jusqu'en 2002, cet enseignement était dispensé dans l'unité de valeur "Techniques avancées pour le multimédia" de la filière ICI. Actuellement, il est traité dans l'unité de valeur "Vision et réalité virtuelle" de la filière I2RV. L'objectif est de former les futurs ingénieurs aux techniques de la reconstruction tridimensionnelle, basées sur la vision stéréoscopique.

On y aborde en particulier :

- La vision stéréoscopique
- Les modèles de caméras
- La mise en correspondance de primitives
- La reconstruction 3D

Filière ICI (département GI), de 2001 – 2002 : 8 h cours, 12 h TD

Filière I2RV (département GI), depuis 2003 : 6 h cours, 14 h TD

### **Compression d'images**

Cet enseignement faisait partie de l'unité de valeur "Codage, cryptographie et compression de données". Il traitait les bases fondamentales de la compression sans perte, la compression d'images fixes et d'images vidéo. Il consistait en :

- La compression sans perte
- La compression JPEG
- La compression MPEG

Filière RT (département GI), de 2001 – 2002 : 8 h cours, 18 h TD, 44 h TP

### **Intelligence artificielle**

L'objectif de cette unité de valeur est d'initier les élèves-ingénieurs aux techniques de la programmation fonctionnelle et la programmation logique utilisées en intelligence artificielle. Le programme est composé des points principaux suivants :

- La représentation des connaissances
- La programmation fonctionnelle
- Les moteur d'inférence et les systèmes à base de connaissances
- La programmation logique

UV de base (département GI), depuis 2001 : 44 h TD

### **Algorithmique et programmation**

C'est une unité de valeur d'initiation à l'algorithmique et à la programmation. Elle est destinée aux élèves-ingénieurs qui n'ont pas suivi le tronc commun de l'UTBM et qui n'ont pas de connaissance en la matière. On y aborde :

- Les types et les structures de contrôle
- Les sous-programmes
- Les pointeurs et les listes chaînées
- Les sous-programmes récursifs

UV de base (département GI), en 2001 : 64 h TD

### **Théorie de la programmation**

Cette unité de valeur présente les outils de construction, transformation et analyse d'algorithmes utiles à la conception de programmes efficaces. Elle traite les points principaux suivants :

- La vérification de programmes
- La complexité des algorithmes
- La transformation d'algorithmes et la récursivité
- Les types abstraits de données

UV de base (département GI), en 2001 : 20 h TD

## **Intelligence artificielle en automatique**

L'objectif de cette unité de valeur est de former les futurs ingénieurs aux nouveaux outils de l'intelligence artificielle pour la commande des systèmes. On traite plus particulièrement :

- Les sous-ensembles flous
- La commande floue
- Les réseaux de neurones
- La commande neuronale

Filière Commande des Systèmes (département GSC), depuis 2002 : 24 h cours, 24 h TD

Master Mécatronique, depuis 2004 : 6 h cours

---

## Annexe 2

### Activités d'encadrement

## Activités d'encadrement

### Encadrement de thèses

#### 1. Mohamed Hariti

Une méthode rapide d'appariement d'images stéréoscopiques – Application à la perception de l'environnement d'un véhicule routier

Thèse soutenue le 20 décembre 2004 à l'Université de Technologie de Belfort-Montbéliard

Co-encadrement à 60% avec A. Koukam

Jury : A. El Moudni (Président), A. Bensrhair (Rapporteur), F. Charpillet (Rapporteur), J. Pallo (Examineur), A. Koukam (Co-directeur de thèse), Y. Ruichek (Co-directeur de thèse)

#### 2. Hazem Issa

Mise en correspondance stéréoscopique par algorithmes génétiques – Nouveaux codages

Thèse soutenue le 19 février 2004 à l'Université des Sciences et Technologies de Lille

Co-encadrement à 50% avec J.-G. Postaire

Jury : O. Colot (Président), A. Bensrhair (Rapporteur), A. Koukam (Rapporteur), J.-C. Burie (Examineur), L. Khoudour (Examineur), J.-G. Postaire (Co-directeur de thèse), Y. Ruichek (Co-directeur de thèse)

#### 3. Tarek Yahiaoui

Etude de faisabilité d'un système de comptage de passagers entrant et sortant d'un autobus

Thèse débutée en septembre 2003 à l'Université des Sciences et Technologies de Lille

Co-encadrement à 40% avec L. Khoudour (40%) et F. Cabestaing (20%)

---

## Encadrement de DEA

### 1. **Olivier Brunner**

Ingénierie des modèles pour les structures et machines intelligentes

DEA Informatique, Automatique et Productique, Université de Technologie de Belfort-Montbéliard, septembre 2004

Co-encadrement à 50% avec L. Thiry

### 2. **Sylvain Boureliou**

Détection et suivi de véhicules par les modèles de contours actifs

DEA Informatique, Automatique et Productique, Université de Technologie de Belfort-Montbéliard, septembre 2004

Encadrement à 100%

### 3. **Laurent Moalic**

Acquisition et visualisation d'informations sur les activités de personnes dans un environnement professionnel

DEA Informatique, Automatique et Productique, Université de Technologie de Belfort-Montbéliard, septembre 2003

Co-encadrement à 50% avec A. Grasso

### 4. **Frédéric Roulland**

Conception et développement d'un prototype d'acquisition du contexte d'un panneau d'affichage dynamique

DEA Informatique, Automatique et Productique, Université de Technologie de Belfort-Montbéliard, septembre 2002

Co-encadrement à 50% avec A. Grasso

## 5. **Hazem Issa**

Mise en correspondance stéréoscopique par algorithmes génétiques

DEA Productique : Automatique et Informatique Industrielle, Université des Sciences et Technologies de Lille, juillet 1999

Co-encadrement à 50% avec J.-G. Postaire

## **Encadrement de stagiaires**

### 1. **Sergio Nogueira**

Navigation autonome d'un véhicule par vision artificielle – Détection d'obstacles et suivi de ligne

Projet de fin d'étude/Master, Université de Technologie de Belfort-Montbéliard, septembre 2005

Encadrement à 100%

### 2. **Slawomir Goryczka**

Localisation d'un véhicule par combinaison des données GPS et visuelles

Master de l'Université des Sciences et Technologies de Cracovie (Pologne), juillet 2005

Encadrement à 100%

### 3. **Szymon Nocon**

Simulation multi-agents d'un véhicule intelligent

Master de l'Université des Sciences et Technologies de Cracovie (Pologne), juillet 2003

Encadrement à 100%

---

## Annexe 3

### Liste des publications

## Publications

### Articles dans des revues scientifiques avec comité de lecture

1. Y. Ruichek. "Multilevel- and neural network-based stereo matching method for real-time obstacle detection using linear cameras", *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, pp. 54–62, 2005.
2. Y. Ruichek, H. Issa, J.G. Postaire. "An evolutionary-based stereo matching method with a multilevel searching strategy", *Soft Computing*, 2005 (à paraître).
3. H. Issa, Y. Ruichek, J.-G. Postaire. "A specific encoding scheme for genetic stereo correspondence searching", *Systemics, Cybernetics and Informatics*, vol. 1, no. 1, 2003.
4. Y. Ruichek, J.-G. Postaire. "A new neural real-time implementation for obstacle detection using linear stereo vision", *Real-Time Imaging*, no. 5, pp. 141–153, 1999.
5. Y. Ruichek, J.-G. Postaire, J.-L. Bruyelle. "A neural approach for obstacle detection with a linear stereoscopic sensor", *Mathematical and Computer Modelling*, vol. 27, no. 9–11, pp. 215–228, 1998.
6. Y. Ruichek, J.-G. Postaire. "A neural matching algorithm for 3D reconstruction from stereo pairs of linear images", *Pattern Recognition Letters*, no. 17, pp. 387–398, 1996.

### Chapitres de livres

1. M. Hariti, Y. Ruichek, A. Koukam. "A voting strategy for matching stereo images", *Book series on Computational Imaging and Vision*, Springer-Verlag Ed., 2005 (à paraître).
2. H. Issa, Y. Ruichek, J.-G. Postaire. "A new genetic encoding for edge point stereo matching", *Advances in Multimedia, Video and Signal Processing Systems*, WSEAS Press, Greece, pp. 169–174, 2002.
3. Y. Ruichek, J.-G. Postaire. "3D environment reconstruction in front of a car with a stereo vision system using a Hopfield neural network", *Road Vehicle Automation II –*

---

Towards systems integration, Edited by Christopher Nwagboso, Wiley Ed., Chichester, UK, pp. 382–391, 1997.

## **Communications dans des conférences internationales avec actes et comités de lecture**

1. S. Nogueira, F. Gechter, Y. Ruichek, A. Koukam, F. Charpillet. "Environment perception for vehicle autonomous navigation in urban areas", accepté à Biennial Workshop on DSP for In-Vehicle and Mobile Systems, Sesimbra (Portugal), septembre 2005.
2. Y. Ruichek, H. Issa, J.G. Postaire, J.C. Burie. "Towards real-time obstacle detection using a hierarchical decomposition methodology for stereo matching with a genetic algorithm", IEEE International Conference on Tools with Artificial Intelligence, pp. 138–147, Florida (Etats-Unis), novembre 2004.
3. M. Hariti, Y. Ruichek, A. Koukam. "A multilevel stereo correspondence searching strategy for real-time obstacle detection using linear cameras", IEEE International Conference on Networking, Sensing and Control, vol. 2, pp. 1002–1007, Taipei (Taiwan), mars 2004.
4. M. Hariti, Y. Ruichek, A. Koukam. "A fast multilevel method for matching stereo images", IEEE International Symposium on Signal Processing and Information Technology, pp. 203–206, Darmstadt (Allemagne), décembre 2003.
5. Y. Ruichek. "A hierarchical neural stereo matching approach for real-time obstacle detection using linear cameras", IEEE International Conference on Intelligent Transportation Systems, pp. 299–304, Shanghai (Chine), octobre 2003.
6. M. Hariti, Y. Ruichek, A. Koukam. "A voting stereo matching method for real-time obstacle detection", IEEE International Conference on Robotics and Automation, vol. 2, pp. 1700–1704, Taipei (Taiwan), septembre 2003.
7. M. Hariti, Y. Ruichek, A. Koukam. "A fast stereo matching method for real-time vehicle front perception with linear cameras", IEEE International Symposium on Intelligent Vehicles, pp. 247–252, Columbus (Etats-Unis), juin 2003.
8. H. Issa, Y. Ruichek, J.G. Postaire. "Stereo correspondence using a genetic scheme

- 
- with a new solution encoding", IEEE International Conference on Systems, Man and Cybernetics, vol. 6, Hammamet (Tunisie), octobre 2002.
9. H. Issa, Y. Ruichek, J.G. Postaire. "Extracting depth information from linear stereo images using a genetic approach", IEEE International Symposium on Intelligent Systems, vol. 1, pp. 285–289, Varna (Bulgarie), septembre 2002.
  10. H. Issa, Y. Ruichek, J.G. Postaire. "A genetic optimization scheme for feature-based stereo matching", World Multiconference on Systemics, Cybernetics and Informatics, vol. 9, Orlando (Etats-Unis), juillet 2002.
  11. Y. Ruichek, H. Issa, J.G. Postaire. "Genetic approach for obstacle detection using linear stereo vision", IEEE International Symposium on Intelligent Vehicles, pp. 261–266, Detroit (Etats-Unis), octobre 2000.
  12. Y. Ruichek, C. Vieren, J.G. Postaire. "Localising moving object using active contour models and a neural stereo matching procedure", International Conference on Engineering Applications of Neural Networks, pp. 355–358, Londres (Grande Bretagne), juin 1996.
  13. P. Vannoorenberghe, Y. Ruichek, J.G. Postaire. "A neural matching procedure for tracking pedestrian in urban environment", International Conference on Engineering Applications of Neural Networks, pp. 347–350, Londres (Grande Bretagne), juin 1996.
  14. Y. Ruichek, J.G. Postaire. "A connexionist approach for the correspondence problem in linear stereo vision", Asian Conference on Computer Vision, vol. 3, pp. 427–431, Singapour, décembre 1995.
  15. Y. Ruichek, J.G. Postaire. "A neural implementation for high speed processing in linear stereo vision", IEEE International Conference on Systems, Man and Cybernetics, pp. 3902–3907, Vancouver (Canada), octobre 1995.
  16. Y. Ruichek, J.G. Postaire. "Real time neural vision for obstacle detection using linear cameras", IEEE International Symposium on Intelligent Vehicles, pp. 524–528, Detroit (Etats-Unis), septembre 1995.
  17. Y. Ruichek, J.G. Postaire. "Extraction of depth information from stereo pairs of linear images using a neural matching algorithm", International Conference on Engineering

---

Applications of Neural Networks, pp. 25–32, Helsinki (Finlande), août 1995.

---

## Communications dans des conférences nationales avec actes et comités de lecture

1. L. Khoudour, T. Yahiaoui, Y. Ruichek. "Développement d'un système de détection de personnes à l'aide de la stéréovision dense – Application au comptage de passagers entrant et sortant d'un autobus", Conférence Internationale Francophone d'Automatique (CIFA), Douz (Tunisie), novembre 2004.
2. M. Hariti, Y. Ruichek, A. Koukam. "Une méthode rapide de mise en correspondance stéréoscopique pour la détection d'obstacles à l'avant d'un véhicule", International Symposium on Technological Innovation for Land Transportation, pp. 279–286, Villeneuve d'Ascq (France), décembre 2003.
3. H. Issa, C. Vieren, Y. Ruichek, L. Khoudour. "Aide à l'exploitation des systèmes de transports guidés par stéréoscopie : une nouvelle approche par algorithme génétique", International Symposium on Technological Innovation for Land Transportation, pp. 305–312, Villeneuve d'Ascq (France), décembre 2003.
4. H. Issa, Y. Ruichek, J.G. Postaire. "A genetic-based stereo matching technique", International Conference on Philosophy and Computer Science – Processes of Evolution in Real and Visual Systems, pp. 101–105, Krakow (Pologne), janvier 2002.
5. Y. Ruichek, J.G. Postaire, L. Macaire, J.C. Burie. "Implantation neuronale pour la mise en correspondance de primitives en stéréovision linéaire", AMSE International Conference on Communication, Signals and Systems, vol. 2, pp. 574–582, Rabat (Maroc), octobre 1995.
6. Y. Ruichek, F. Cabestaing, C. Vieren. "Contribution d'une approche neuromimétique à l'appariement stéréoscopique des contours actifs", JNAEETSA, pp. 320–325, Marrakech (Maroc), octobre 1996.
7. Y. Ruichek, J.G. Postaire. "Une approche neuronale pour la stéréovision linéaire", Colloque sur le Traitement du Signal et des Images (GRETSI), pp. 889–892, Juan les Pins (France), septembre 1995.

## **Diffusion de la connaissance avec actes**

1. Y. Ruichek, F. Gechter, A. Koukam, F. Charpillet. "Véhicules Intelligents – Perception de l'environnement et Navigation", Colloque sur les Systèmes Embarqués et Transports Terrestres, Belfort, mars 2005.
2. Y. Ruichek, M. Hariti, A. Koukam. "Système d'aide à la conduite – Perception de l'environnement d'un véhicule par anté-vision et rétro-vision stéréoscopique", Journée Technique sur la Sécurité dans les Transports Terrestres, Belfort, mai 2002.

## **Rapports**

1. Y. Ruichek, M. Hariti, A. Koukam. "Système d'aide à la conduite par combinaison de la perception des environnements avant et arrière d'un véhicule", Rapport de fin de contrat "aide aux jeunes chercheurs pour l'innovation" (ANVAR 2002 – 2003 : No. J 02.04.017I JJ), avril 2003.

## **Articles soumis**

1. Y. Ruichek, M. Hariti, A. Koukam. "A score-based stereo matching algorithm for high speed obstacle detection", Soumis à International Journal of Image and Vision Computing, février 2005.

---

## BIBLIOGRAPHIE

- [1] M. Accame and F. De Natale. *Disparity estimation for stereo sequences based on adaptive size hierarchical block matching*. Proceedings of the 8th International Conference on Image Analysis and Processing, pages 453–458, San Remo, Italie, 1995.
- [2] P. Anandon. A computational framework and an algorithm for the measurement of visual motion. *International Journal of Computer Vision*, 2(3) :283–310, 1989.
- [3] R. Aufrère, R. Chapuis, and F. Chausse. *A fast and robust vision based road following algorithm*. Proceedings of the IEEE Intelligent Vehicles Symposium, pages 192–197, Dearborn, Etats-Unis, octobre 2000.
- [4] N. Ayache. *Vision stéréoscopique et perception multisensorielle – Application à la robotique mobile*. InterEditions, Paris, France, 1989.
- [5] N. Ayache and B. Faferjon. *Un algorithme rapide de stéréoscopie passive utilisant la prédiction et la vérification récursive d’hypothèses*. Proceedings du 5ème Congrès AFCET sur la Reconnaissance des Formes et Intelligence Artificielle, pages 1217–1232, Grenoble, France, novembre 1985.
- [6] H.H. Baker and T.O. Binford. *Depth from edge and intensity based stereo*. Proceedings of the 7th International Joint Conference on Artificial Intelligence, pages 631–636, Vancouver, Canada, août 1981.

- 
- [7] D.H. Ballard and C.M. Brown. *Computer vision*. Prentice Hall Inc., Englewood Cliffs, New Jersey, Etats-Unis, 1982.
- [8] S.T. Barnard. Stochastic stereo matching over scale. *International Journal of Computer Vision*, 3(1) :17–32, 1989.
- [9] T. Bücher, C.C. Edelbrunner, C. Igel, D. Kastrup, I. Leefken, G. Lorenz, A. Steinhage, and W. Von Seelen. Image processing and behaviour planning for intelligent vehicles. *IEEE Trans. on Industrial Electronics*, 90(1) :62–75, février 2003.
- [10] N. Benamrane. *Contribution à la vision stéréoscopique par mise en correspondance de régions*. Thèse de Doctorat, Université de Valenciennes et du Hainaut Cambresis, Valenciennes, France, 1994.
- [11] M. Benard. Automatic stereophotogrammetry : a method based on feature detection and dynamic programming. *Photogrammetria*, (39) :169–181, 1984.
- [12] A. Bensrhair, P. Miché, and R. Debrie. Fast and automatic stereo vision matching algorithm based on dynamic programming method. *Pattern Recognition Letters*, 17(5) :457–466, mai 1996.
- [13] M. Bertozzi, A. Broggi, and A. Fascioli. *Development and test of an intelligent vehicle prototype*. Proceedings of the 7th World Congress on Intelligent Transportation Systems, novembre 2000.
- [14] S. Birchfield and C. Tomasi. A pixel dissimilarity measure that is insensitive to image sampling. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(4) :401–406, avril 1998.
- [15] A.F. Bobick and S.S. Intille. Large occlusion stereo. *International Journal of Computer Vision*, 33(3) :181–200, 1999.
- [16] A.C. Bovik, N. Gopal, T. Emmoth, and A. Restrepo. Localized measurement of emergent image frequencies by gabor wavelets. *IEEE Trans. on Information Theory*, 38(2) :691–712, 1992.
- [17] Y. Boykov, O. Veksler, and R. Zabih. A variable window approach to early vision. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(12) :1283–1294, décembre 1998.

- 
- [18] M.Z. Brown, D. Burschka, and G.D. Hager. Advances in computational stereo. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(8) :993–1008, août 2003.
- [19] J.L. Bruyelle. *Conception et réalisation d'un dispositif de prise de vue stéréoscopique linéaire – Application à la détection d'obstacles à l'avant des véhicules routiers*. Thèse de Doctorat, Université des Sciences et Technologies de Lille, Lille, France, décembre 1994.
- [20] J.M.H. Du Buf, M. Kardan, and M. Spann. Texture feature performance for image segmentation. *Pattern Recognition*, 23(3–4) :291–309, 1990.
- [21] J.C. Burie. *Mise en correspondance d'images linéaires stéréoscopiques – Application à la détection d'obstacles à l'avant des véhicules routiers*. Thèse de Doctorat, Université des Sciences et Technologies de Lille, Lille, France, septembre 1995.
- [22] J.F. Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pages 679–698, novembre 1986.
- [23] S. Castan and J. shen. *A stereo algorithm taking into account the perspective distortions*. Proceedings of the International Conference on Pattern Recognition, pages 441–443, Montréal, Canada, 1984.
- [24] R. Chapuis. *Localisation et suivi de route pour l'aide à la conduite*. Habilitation à Diriger des Recherches, Université Blaise-Pascal, Clermont-Ferrand, France, décembre 2000.
- [25] B. Chebaro, L. Massipailles, and S. Castan. *Description polygonale basée sur l'extraction d'îlots de confiance*. Proceedings du 8ème Congrès sur la Reconnaissance des Formes et Intelligence Artificielle, volume 2, pages 841–846, Lyon, France, novembre 1991.
- [26] L. Cheng and T. Caelli. *Bayesian stereo matching*. Proceedings of the International Conference on Computer Vision and pattern Recognition, volume 12, pages 192–199, Washington, Etats-Unis, juin 2004.
- [27] W.J. Christmas, J. Kittler, and M. Petrou. Structural matching in computer vi-

- 
- sion using probabilistic relaxation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(8) :749–764, août 1995.
- [28] J.P. Cocquerez and S. Philipp. *Analyse d'images : filtrage et segmentation*. Edition Masson, Paris, France, 1995.
- [29] J.P. Cocquerez and O. Monga. *Matching regions in stereovision*. Proceedings of the 5th Scandinavian Conference on Image Analysis, pages 295–308, Stockholm, Suède, 1987.
- [30] L.D. Cohen. On active contour models and balloon. *Computer Vision, Graphics and Image Processing : Image Understanding*, 53(2) :211–218, mars 1991.
- [31] O. Colle. *Vision stéréoscopique à l'aide de deux caméras linéaires – Application à la robotique mobile*. Thèse de Doctorat, Institut National des Sciences Appliquées de Lyon, Lyon, France, novembre 1990.
- [32] I.J. Cox and al. A maximum likelihood stereo algorithm. *Computer Vision and Image Understanding*, 63(3) :543–567, 1996.
- [33] A. Crouzil. *Perception du relief et du mouvement par analyse d'une séquence stéréoscopique d'images*. Thèse de Doctorat, Université Paul Sabatier, Toulouse, France, 1997.
- [34] J.M. Cruz, G. Pajares, and J. Aranda. A neural network model in stereovision matching. *Neural Networks*, 8(5) :805–813, 1995.
- [35] J. DaSilva, P. Simoni, and K. Bharadwaj. *Multiple correspondence in stereo vision under a genetic algorithm approach*. Proceedings of the 8th Brazilian Symposium on Computer Graphics and Image Processing, pages 52–59, 2000.
- [36] L.S. Davis and A. Rosenfield. Cooperating process for low-level vision : A survey. *Artificial Intelligence*, 17(1–3) :245–263, août 1981.
- [37] A. Robert de Saint-Vincent. *Perception et modélisation de l'environnement d'un robot mobile : une approche par stéréovision*. Thèse de Doctorat, Université Paul Sabatier, Toulouse, France, 1986.
- [38] P. Deloof, N. Hease, and S. Baranowski. Recherche bibliographique sur les radars de prévention des collisions. *Recherche Transports Sécurité*, (26) :15–22, juin 1990.

- 
- [39] P. Deloof, P.A. Rolland, and N. Hease. La prévention des collisions routières par radar anticollision micro-onde. Rapport 25, Institut National de Recherche sur les Transports et leur Sécurité, juillet 1994.
- [40] R. Deriche. Fast algorithms for low-level vision. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(1) :78–87, 1990.
- [41] E.D. Dickmanns. *The development of machine vision for road vehicles in the last decade*. Proceedings of the IEEE Intelligent Vehicles Symposium, pages 268–281, Versailles, France, juin 2002.
- [42] D. Dooze. *Conception et réalisation d'un stéréoscope bimodal à portée variable – Application à la détection d'obstacles à l'avant de véhicules guidés automatisés*. Thèse de Doctorat, Université des Sciences et Technologies de Lille, Lille, France, juin 2001.
- [43] A. Ewald and V. Willhoeft. *Laser scanners for obstacle detection in automotive application*. Proceedings of the IEEE Intelligent Vehicles Symposium, pages 682–687, Dearborn, Etats-Unis, octobre 2000.
- [44] Y. Fang, I. Masaki, and B. Horn. Depth-based target segmentation for intelligent vehicles : Fusion of radar and binocular stereo. *IEEE Trans. on Intelligent Transportation Systems*, 3(3) :196–202, septembre 2002.
- [45] O.D. Faugeras. Quelques pas vers la vision artificielle en trois dimensions. *Techniques et Sciences Informatiques*, 7(6) :547–590, 1988.
- [46] O.D. Faugeras, B. Hotz, H. Mathieu, T. Viéville, Z. Zhang, P. Fua, E. Théron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, and C. Proy. Real-time correlation-based stereo : algorithms, implementations and applications. Rapport de recherche RR 2013, Institut National de Recherche en Informatique et en Automatique, Sophia-Antipolis, France, août 1993.
- [47] D. Fleet, A. Jepson, and M. Jenkin. Phase-based disparity measurement. *CVGIP : Image Understanding*, 53 :198–210, 1991.
- [48] D.J. Fleet. *Disparity from local weighted phase-correlation*. Proceedings of the IEEE

- 
- International Conference on Systems, Man and Cybernetics, pages 48–56, octobre 1994.
- [49] M. Fontaine. *Segmentation non supervisée d'images couleur par analyse de la connexité des pixels*. Thèse de Doctorat, Université des Sciences et Technologies de Lille, Lille, France, 2001.
- [50] N. Funabiki, Y. Takefuji, and K.C. Lee. Comparisons of seven neural network models on traffic control problems in multistage interconnection networks. *IEEE Trans. on Computers*, 42(4) :497–501, avril 1993.
- [51] A. Fusiello, V. Roberto, and E. Trucco. *Efficient stereo with multiple windowing*. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, pages 858–863, San Juan, Puerto Rico, juin 1997.
- [52] A. Fusiello, E. Trucco, and A. Verri. A compact algorithm for rectification of stereo pairs. *Machine Vision and Applications*, 12(1) :16–22, 2000.
- [53] A.D. Galway, H. Knutsson, and R. Wilson. *Multiresolution estimation of 2D disparity using a frequency domain approach*. Proceedings of the British Machine Vision Conference, pages 227–236, septembre 1992.
- [54] D. Geiger, B. Ladendorf, and A. Yuille. Occlusions and binocular stereo. *Lecture Notes in Computer Science*, 588 :425–433, 1992.
- [55] A. Gern, U. Franke, and P. Levi. *Advanced lane recognition fusing vision and radar*. Proceedings of the IEEE Intelligent Vehicles Symposium, pages 45–51, Dearborn, Etats-Unis, octobre 2000.
- [56] B. Gai-Checa, P. Bouthémy, and T. Vieville. Détection d'objets en mouvement. Rapport Technique RR 1906, Institut National de Recherche en Informatique et en Automatique, Sophia-Antipolis, juin 1993.
- [57] W.R. Gilks, S. Richardson, and D.J. Soieghalter. *Marcov chain Monte Carlo in practice*. Chapman and Hall, London, 1996.
- [58] F. Glover and al. *modern heuristic techniques for combinatorial problems*. Black-Well Scientific Publications. Oxford, UK, 1993.

- 
- [59] D. Goldberg. *Genetic algorithms in search, optimization and machine learning*. Edition Addison-Wesley, Reading, MA, Etats-Unis, 1989.
- [60] M. Gong and Y. Yang. Genetic-based stereo algorithm and disparity map evaluation. *International Journal of Computer Vision*, 47(1) :63–77, 2002.
- [61] J.Y. Goulerams and P. Liatis. Hybrid symbiotic genetic optimisation for robust edge-based stereo correspondence. *Pattern Recognition*, 34(12) :2477–2496, 2001.
- [62] W.E.L. Grimson. *From images to surfaces*. MIT Press, Cambridge, Etats-Unis, 1981.
- [63] E.R. Hancock and J. Kittler. Discrete relaxation. *Pattern Recognition*, 23(7) :711–733, 1990.
- [64] J. Hancock, M. Hebert, and C. Thorpe. *Laser intensity-based obstacle detection*. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, volume 3, pages 1541–1546, octobre 1998.
- [65] R. Hartley. Theory and practice of projective rectification. *International Journal of Computer Vision*, 35(2) :115–127, 1999.
- [66] S. Hattori, A. Okamoto, and H. Hasegawa. *Stereo matching by simulated annealing incorporating a diffusion equation*. Proceedings of ASPRS Annual Conference, pages 1030–1041, 1998.
- [67] J. Heikkila and O. Silven. *A four-step camera calibration procedure with implicit image correction*. Proceedings of the International Conference on Computer Vision and Pattern Recognition, volume 1, pages 1106–1112, San Juan, Etats-Unis, 1997.
- [68] E.E. Hemayed, A. Sandbek, A.G. Wassal, and A.A. Farag. *Investigation of stereo-based 3D surface reconstruction*. Proceedings of SPIE, volume 3023, pages 191–2002, mars 1997.
- [69] J.J. Hopfield and D.W. Tank. Neural computation of decisions in optimization problems. *Biological Cybernetics*, (52) :141–152, 1985.
- [70] R. Horaud and O. Monga. *Structural matching for stereo vision*. Proceedings of the International Conference on Pattern Recognition, Rome, Italie, novembre 1988.

- 
- [71] R. Horaud and T. Skordas. Stereo matching through feature grouping and maximal cliques. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(11) :1168–1180, novembre 1989.
- [72] R. Horaud and T. Skordas. *Vision par ordinateur, outils fondamentaux*. Traité des nouvelles technologies, série Informatique. Edition Hermès, Paris, France, 1993.
- [73] L. Héroult. *Réseaux de neurones récurrents pour l'optimisation combinatoire – Application à la théorie des graphes et à la vision par ordinateur*. Thèse de Doctorat, Institut National Polytechnique de Grenoble, Grenoble, France, février 1991.
- [74] J.E. Hu and P. Siy. *An ordering-oriented Hopfield network and its application in stereo vision*. Proceedings of the SPIE International Conference on Application of Artificial Neural Networks IV, volume 1965, pages 556–567, Orlando, Etats-Unis, avril 1993.
- [75] J.E. Hu and P. Siy. *Stereo correspondence through multiple constraint neural network*. Proceedings of the IEEE International Conference on Neural Networks, pages 126–131, San Francisco, Etats-Unis, 1993.
- [76] J.S. Huang and H.C. Liu. Stereo vision using a microcanonical mean field annealing neural network. *Network : Computation in Neural Systems*, 8(1) :87–104, 1997.
- [77] D. Hubel. The brain. *Scientific American*, 241(3) :39–47, septembre 1979.
- [78] R.M. Inigo and T. Tkacik. *Mobile robot operation in real-time with linear image array based vision*. Proceedings of the IEEE International Intelligent Control Symposium, pages 228–233, Philadelphia, Etats-Unis, janvier 1987.
- [79] R.M. Inigo, T. Tkacik, and E.S. McVey. *The application of linear image arrays to mobile robot guidance and navigation*. Proceedings of the International Conference on Automated Guided Vehicle Systems, pages 157–168, Stockholm, Suède, octobre 1985.
- [80] A. Jagmohan, M. Singh, and N. Ahuja. *Dense stereo matching using kernel maximum likelihood estimation*. Proceedings of the 17th International Conference on Pattern Recognition, volume 3, pages 28–31, août 2004.

- 
- [81] A.D. Jepson and M.R.M. Jenkin. *The fast computation of disparity from phase differences*. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, pages 398–403, San Diego, Etats-Unis, juin 1989.
- [82] T. Kanade and M. Okutomi. A stereo matching algorithm with an adaptive window : theory and experiment. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(9) :920–932, septembre 1994.
- [83] M. Kass, A. Witkin, and D. Terzopoulos. Snakes : Active contour models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 3 :259–267, 1987.
- [84] T. Kato, Y. Ninomiya, and I. Masaki. An obstacle detection method by fusion of radar and motion stereo. *IEEE Trans. on Intelligent Transportation Systems*, 3(3) :182–188, septembre 2002.
- [85] A. Khotanzad, A. Bokil, and Y.W. Lee. Stereopsis by constraint learning feed-forward neural networks. *IEEE Trans. on Neural Networks*, 4(2) :332–342, mars 1993.
- [86] A. Khotanzad and Y.W. Lee. *Depth perception by a neural network*. Proceedings of the Midcon/90 Conference Record Ventura, pages 424–427, Dallas, Etats-Unis, septembre 1990.
- [87] A. Khotanzad and Y.W. Lee. *Stereopsis by a neural network which learns the constraints*. Advances in Neural Information Processing Systems, Morgan Kaufmann Ed., volume 3, pages 327–334, 1991.
- [88] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598) :671–680, 1983.
- [89] L. Kiyong and L. Joonwoong. *Generic obstacle detection on roads by dynamic programming for remapped stereo images to an overhead view*. Proceedings of the IEEE International Conference on Networking, Sensing and Control, pages 897–902, Taipei, Taiwan, 2004.
- [90] A. Koschan, V. Rodehorst, and K. Spiller. *Color stereo vision using hierarchical block matching and active color illumination*. Proceedings of the 13th International

- 
- Conference on Pattern Recognition, volume 1, pages 835–839, Viennes, Autriche, août 1996.
- [91] E. Krotkov, K. Henriksen, and R. Kories. Stereo ranging with verging cameras. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(12) :1200–1205, décembre 1990.
- [92] P.M.J.V. Laarhoven and E.H.L. Aarts. *Simulated annealing : Theory and applications*. Norwell, MA : Kluwer, 1989.
- [93] R. Labayrade, D. Aubert, and J. P. Tarel. *Real time obstacle detection on non-flat road geometry through V-disparity representation*. Proceedings of the IEEE Intelligent Vehicles Symposium, Versailles, France, juin 2002.
- [94] J. Laneurit, R. Chapuis, and F. Chausse. Accurate vehicle positioning on a numerical map. *International Journal of Control, Automation and Systems*, 3(1) :15–31, mars 2005.
- [95] D. Langer and C. Thorpe. *Sonar based outdoor vehicle navigation and collision avoidance*. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, volume 2, pages 1445–1450, juillet 1992.
- [96] P. Lasserre. *Vision pour la robotique mobile en environnement naturel*. Thèse de Doctorat, Université Paul Sabatier, Toulouse, France, 1996.
- [97] H.C. Lee and D.R. Cok. Detecting boundaries in a vector field. *IEEE Trans. on Signal Processing*, 39(5) :1181–1194, mai 1991.
- [98] J.J. Lee, J.C. Shim, and Y.H. Ha. Stereo correspondence using the Hopfield neural network of a new energy function. *Pattern Recognition*, 27(11) :1513–1522, 1994.
- [99] S.H. Lee and J.J. Leou. A dynamic programming approach to line segment matching in stereo vision. *Pattern Recognition*, 27(8) :961–986, août 1994.
- [100] M.S. Lew, K.W. Wong, and T.S. Huang. *Multi-scale stereo matching*. Proceedings of the 11th IAPR International Conference on Pattern Recognition, volume 1, pages 620–623, La Hague, Pays-Bas, 1992.
- [101] F. Leymarie and M.D. Levine. Tracking deformable objects in the plane using an

- 
- active contour model. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(1) :56–75, 1993.
- [102] P. Long-Limozin. *Vision stéréoscopique appliquée à la robotique*. Thèse de Doctorat, Université de Nice, France, 1986.
- [103] J.L. Lotti and G. Giraudon. *Adaptive window algorithm for aerial image stereo*. Proceedings of the 12th IAPR International Conference on Pattern Recognition, volume 1, pages 701–703, octobre 1994.
- [104] J.L. Lotti and G. Giraudon. *Correlation algorithm with adaptive window for aerial image in stereo vision*. Proceedings of the SPIE Image and Signal Processing for Remote Sensing, volume 2315, pages 76–87, Rome, Italie, 1994.
- [105] J. Louchet. *Stereo analysis using individual evolution strategy*. Proceedings of the International Conference on Pattern Recognition, pages 170–176, 2000.
- [106] D. Marr and T. Poggio. Cooperative computation of stereo disparity. *Science*, (194) :283–287, 1976.
- [107] F. Martinez, B. Thiesse, and P. Marthon. *Reconstruction de scènes 3D en vision binoculaire animée*. 4ème Journées ORASIS, GDR-PRC Communication Homme–Machine, pages 13–14, Mulhouse, France, octobre 1993.
- [108] S. Möckel, F. Scherer, and P.F. Schuster. *Multi-sensor obstacle detection on railway tracks*. Proceedings of the IEEE Intelligent Vehicles Symposium, pages 42–46, Columbus, Etats-Unis, juin 2003.
- [109] G. Medioni and R.N. Nevatia. Matching images using linear features. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6(6) :675–685, 1984.
- [110] G. Medioni and R.N. Nevatia. Segment-based stereo matching. *Computer Vision, Graphics and Image Processing*, 31 :2–18, 1985.
- [111] P. Miché and R. Debrie. Fast and self-adaptive image segmentation using extended declivity. *Annals of Telecommunication*, 50(3–4) :401–410, 1995.
- [112] Z. Michalewicz. *Genetic algorithms + Data structure = Evolution programs*. Springer-Verlag, Berlin, Allemagne, seconde édition edition, 1994.

- 
- [113] H.P. Moravec. *Rover visual obstacle avoidance*. Proceedings of the International Joint Conference on Artificial Intelligence, pages 785–790, Vancouver, Canada, août 1981.
- [114] M.S. Mousavi and R.J. Schalkoff. *A neural network approach for stereo vision*. Proceedings of the IEEE SouthEastConf90, pages 808–812, 1990.
- [115] M.S. Mousavi and R.J. Schalkoff. A neural network implementation of stereo vision using a multilayer feedback architecture. *IEEE Trans. on Systems, Man and Cybernetics*, 24(8) :1220–1238, août 1994.
- [116] J. Mulligan and K. Daniilidis. Predicting disparity windows for real-time stereo. *Lecture Notes in Computer Science*, 1842 :220–235, 2000.
- [117] N.M. Nasrabadi. A stereo vision technique using curve-segments and relaxation matching. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(5) :566–572, mai 1992.
- [118] N.M. Nasrabadi and C.Y. Choo. Hopfield network for stereo vision correspondence. *IEEE Trans. on Neural Networks*, 3(1) :5–13, janvier 1992.
- [119] R. Nevatia and K. Babu. Linear feature extraction and detection. *Computer Vision, Graphics and Image Processing*, 13 :257–269, 1980.
- [120] S. Nichani. *Solving the correspondence problem using a Hopfield network*. Proceedings of the IEEE International Conference on Neural Networks, volume 6, pages 4107–4112, Orlando, Etats-Unis, 1994.
- [121] H.K. Nishihara. Practical real-time imaging stereo matcher. *Optical Engineering*, 23(5) :536–545, 1984.
- [122] J.A. Noble. Finding corners. *Image and Vision Computing*, (6) :121–128, 1988.
- [123] J. Ohm, K. Gruneberg, E. Hendriks, E. Izquierdo, D. Kalivas, M. Karl, and D. Papadimitos. A real-time hardware system for stereoscopic videoconferencing with viewpoint adaptation. *Image Communication*, 14(1) :147–171, 1998.
- [124] Y. Ohta and T. Kanade. Stereo by intra- and inter-scanline search using dynamic programming. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 7(2) :139–154, mars 1985.

- 
- [125] N. Okutomi and T. Kanade. A locally adaptive window for signal matching. *International Journal of Computer Vision*, 7(2) :143–162, 1992.
- [126] M. Ouali, C. Lourceau, and D. Ziou. *Dense disparity estimation using gabor filters and image derivatives*. Proceedings of the 2nd International Conference on 3D Imaging and Modeling, volume 1, pages 483–490, Ottawa, Canada, 1999.
- [127] G. Pajares, J.M. Cruz, and J. Aranda. Relaxation by Hopfield network in stereo image matching. *Pattern Recognition*, 31(5) :561–574, 1998.
- [128] G. Pajares, J.M. Cruz, and J. Aranda. Stereo matching based on the self-organizing feature-mapping algorithm. *Pattern Recognition Letters*, 19(3–4) :319–330, 1998.
- [129] G. Pajares, J.M. Cruz, and J.A. Lopez-Orozco. Relaxation labeling in stereo image matching. *Pattern Recognition*, 33(1) :53–68, 2000.
- [130] G. Pajares and J.M. de la Cruz. On combining support vector machines and simulated annealing in stereovision matching. *IEEE Trans. on Systems, Man and Cybernetics – Part B : Cybernetics*, 34(4) :1646–1657, août 2004.
- [131] N. Paragios and R. Deriche. Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(3) :266–280, mars 2000.
- [132] M. Perez. *Stéréovision dense par traitement adaptatif temps réel : Algorithmes et implantation*. Thèse de Doctorat, Université des Sciences et Technologies de Lille, Lille, France, 2005.
- [133] C.B.U. Perwass and G. Sommer. *An iterative Bayesian technique for dense image point matching*. Dynamic Perception, IOS Press infix Ed., pages 283–288, 2003.
- [134] S.B. Pollard and al. *Disparity gradient, lipschitz continuity and computing binocular correspondence*. Proceedings of the 3rd International Symposium on Robot Research, pages 19–26, France, 1986.
- [135] S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby. Pmf : A stereo correspondence algorithm using a disparity gradient limit. *Perception*, (14) :449–470, 1985.
- [136] K.E. Price. Relaxation matching techniques – a comparison. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 7(5) :617–623, 1985.

- 
- [137] T. Quiguer, P. Miché, and R. Debrie. *Segmentation method by self-adaptive thresholding*. Proceedings of the Conference on Image Analysis and Processing, pages 34–41, Como, Italie, 1991.
- [138] R. Ronchard. Region-based strategies for active contour models. *Computer Vision*, 13(2) :229–251, octobre 1994.
- [139] Y. Ruichek. *Stéréovision linéaire par réseau de neurones de Hopfield : Application à la détection d’obstacles à l’avant des véhicules routiers*. Thèse de Doctorat, Université des Sciences et Technologies de Lille, Lille, France, 1997.
- [140] Y. Ruichek and J.G. Postaire. A neural matching algorithm for 3D reconstruction from stereo pairs of linear images. *Pattern Recognition Letters*, 17(4) :387–398, 1996.
- [141] Y. Ruichek and J.G. Postaire. A new neural real-time implementation for obstacle detection using linear stereo vision. *Real-Time Imaging*, (5) :141–153, 1999.
- [142] Y. Ruichek, J.G. Postaire, and J.L. Bruyelle. A neural approach for obstacle detection with a linear stereoscopic sensor. *Mathematical and Computer Modelling*, 27(9–11) :215–228, 1998.
- [143] Y. Ruichek, C. Vieren, and J.G. Postaire. *Localising moving objects using active contour models and a neural stereo matching procedure*. Proceedings of the International Conference on Engineering Applications of Neural Networks, pages 355–358, Londres, Grande Bretagne, juin 1996.
- [144] H. Saito and M. Mori. Application of genetic algorithms to stereo matching of images. *Pattern Recognition Letters*, 16(8) :815–821, 1995.
- [145] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Computer Vision*, 47(1) :7–42, 2002.
- [146] M. Selsis. *Application des modèles de contours actifs au suivi et à la localisation 3D d’objets en mouvement*. Thèse de Doctorat, Université des Sciences et Technologies de Lille, Lille, France, 1996.
- [147] N. Sgouros, G. Papakonstantinou, and P. Tsanakas. *Localized qualitative Navigation for indoor for indoor environments*. Proceedings of the IEEE International

- 
- Conference on Robotics and Automation, pages 921–926, Minneapolis, Etats-Unis, avril 1996.
- [148] J. Shen, S. Castan, and J. Zhao. *Stereo vision by pyramidal BLI graph matching*. Proceedings of SPIE'87, Cannes, France, 1987.
- [149] M.K. Singh. *Image segmentation and robust estimation using Parzen windows*. PhD Thesis, University of Illinois, Etats-Unis, 2003.
- [150] T. Skordas. *Mise en correspondance et reconstruction stéréo utilisant une description structurelle des images*. Thèse de Doctorat, Institut National Polytechnique de Grenoble, France, 1988.
- [151] T. Skordas and R. Horaud. Mise en correspondance structurelle pour la vision stéréoscopique. *Techniques et Sciences Informatiques*, 7(6) :591–608, 1988.
- [152] C. Stiller, W. Pöschmüller, and B. Hürtgen. *Stereo vision in driver assistance systems*. Proceedings of the IEEE International Conference on Intelligent Transportation Systems, pages 888–893, Boston, Etats-Unis, novembre 1997.
- [153] J. Sun, N.N. Zheng, and H.Y. Shum. Stereo matching using belief propagation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(7) :787–800, juillet 2003.
- [154] C. Thorpe. *Vision and navigation – The Carnegie Mellon NavLab*. Kluwer Academic Publishers, Boston, Etats-Unis, 1990.
- [155] F.C. Tien. Solving line-feature stereo matching with genetic algorithms in Hough space. *Journal of the Chinese Institute of Industrial Engineers*, 21(5) :516–526, 2004.
- [156] F.C. Tien and C.A. Chang. Neural network for precise 3D measurement in stereo vision system. *International Journal of Production Research*, 37(9) :1935–1948, 1999.
- [157] S. Tokoro. *Automotive application systems of a millimeter-wave radar*. Proceedings of the IEEE International Intelligent Vehicles Symposium, pages 260–265, Tokyo, Japon, septembre 1996.

- 
- [158] R. Tsai. Multiframe image point matching and 3D surface reconstruction. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 5(2) :159–174, 1986.
- [159] R.Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf *tv* cameras and lenses. *IEEE Trans. on Robotics and Automation*, 3(4) :p. 323–344, 1987.
- [160] V. Ultré. *Contribution à la segmentation d'images de mosaïque couleur*. Thèse de Doctorat, Université des Sciences et Technologies de Lille, Lille, France, 1996.
- [161] R. Vaillant and L. Gueguen. *Genetic algorithms applied to binocular stereovision*. Proceedings of the European Conference in Computer Vision, pages 193–198, 1994.
- [162] N. Vandenbrouke. *Segmentation d'images couleur par classification de pixels dans des espaces d'attributs colorimétriques adaptés. Application à l'analyse d'images de football*. Thèse de Doctorat, Université des Sciences et Technologies de Lille, Lille, France, 2000.
- [163] P. Vannoorenberghe. *Détection de mouvement par analyse de séquences d'images monoculaires – Application à l'estimation de flux de piétons en milieu urbain*. Thèse de Doctorat, Université du Littoral, Calais, France, juillet 1997.
- [164] P. Vannoorenberghe, Y. Ruichek, and J.G. Postaire. *A neural matching procedure for tracking pedestrian in urban environment*. Proceedings of the International Conference on Engineering Applications of Neural Networks, pages 347–350, Londres, Grande Bretagne, juin 1996.
- [165] L. Vinet and A. Gagalowicz. *Mise en correspondance de régions utilisant la cohérence de leurs voisinage*. Proceedings du 8ème Congrès de Reconnaissance des Formes et Intelligence Artificielle, volume 2, pages 821–826, Lyon, France, novembre 1991.
- [166] Y.F. Wan. *Un opérateur hyperbolique pour la détection de contours dans les images numériques*. Thèse de Doctorat, Université des Sciences et Technologies de Lille, Lille, France, 1996.
- [167] C.C. Wang, C. Thorpe, and A. Suppe. *Ladar-based detection and tracking of moving*

- 
- objects from a ground vehicle at high speeds*. Proceedings of the IEEE International Intelligent Vehicles Symposium, pages 416–421, Columbus, Etats-Unis, juin 2003.
- [168] J.H. Wang and C.P. Hsiao. Representation-burden conservation network applied to learning *vsq*. *Neural Processing Letters*, 5(3) :209–217, 1997.
- [169] J.H. Wang and C.P. Hsiao. On disparity matching in stereo vision via a neural network framework. *Proc. National Science Council ROC(A)*, 23(5) :665–678, 1999.
- [170] J. Weng. Image matching using windowed fourier phase. *International Journal of Computer Vision*, 11 :211–236, 1994.
- [171] R. Williamson and C. Thorpe. *A trinocular stereo system for highway obstacle detection*. Proceedings of the IEEE International Conference on Robotics and Automation, volume 3, pages 2267–2273, Detroit, Etats-Unis, mai 1999.
- [172] B. Worobel-Dautcourt. *Perception de la distance par mise en correspondance de régions entre des images stéréoscopiques*. Thèse de Doctorat, Institut National Polytechnique de Lorraine, Nancy, France, 1988.
- [173] C. Xu and J.L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Trans. on Image Processing*, 7(3) :359–369, mars 1998.
- [174] J. Yedidia, W.T. Freeman, and Y. Weiss. *Understanding belief propagation and its generalizations*. Exploring Artificial Intelligence in the New Millennium, Gerhard Lakemeyer Ed., pages 239–269, 2003.
- [175] R.K.K. Yip and W.P. Ho. A multi-level dynamic programming method for stereo line matching. *Pattern Recognition Letters*, 19(9) :839–855, juillet 1998.
- [176] N. Yokoya. *Surface reconstruction directly from binocular stereo images by multiscale-multistage regularization*. Proceedings of the 11th IAPR International Conference on Pattern Recognition, pages 642–646, La Hague, Pays-Bas, 1992.
- [177] P. Zhan, D.J. Lee, and R. Beard. *Solving correspondence problem with 1D signal matching*. Proceedings of SPIE, volume 5608, pages 207–217, octobre 2004.
- [178] L. Zhang and S.M. Seitz. *Parameter estimation for MRF stereo*. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, San Diego, Etats-Unis, juin 2005.

- [179] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(11) :1330–1334, 2000.
- [180] J. Zhao. *Extraction d'information tridimensionnelle par stéréovision*. Thèse de Doctorat, Université Paul Sabatier, Toulouse, France, 1989.
- [181] Q. Zheng and R. Chellappa. Automatic feature point extraction and tracking in image sequences for arbitrary camera motion. *International Journal of Computer Vision*, 15(1–2) :31–76, 1995.