

UNIVERSITÉ DES SCIENCES ET TECHNOLOGIES DE LILLE 1
U.F.R. d'Informatique, Électronique, Électrotechnique et Automatique
Laboratoire d'Automatique I³D

Numéro attribué par la bibliothèque : 3441

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE LILLE 1

Discipline : Automatique et Informatique Industrielle

Présentée et soutenue publiquement

par

Hazem ISSA

le 19 Février 2004

***MISE EN CORRESPONDANCE STÉRÉOSCOPIQUE PAR
ALGORITHMES GÉNÉTIQUES : NOUVEAUX CODAGES***

JURY :

M. Olivier COLOT	Professeur à l'Université des Sciences et Technologies de Lille, Président de Jury
M. Abderrafiaa KOUKAM	Professeur à l'Université de Technologie de Belfort-Montbéliard, Rapporteur
M. Abdelaziz BENSRAIR	Professeur à l'Université de Rouen, Rapporteur
M. Jean-Christophe BURIE	Maître de Conférences à l'Université de la Rochelle, Examineur
M. Louahdi KHOUDOUR	Ingénieur de Recherche à l'INRETS - LEOST, Examineur
M. Yassine RUICHEK	Maître de Conférences à l'Université de Technologie de Belfort-Montbéliard, Co-Directeur de Recherche
M. Jack-Gérard POSTAIRE	Professeur à l'Université des Sciences et Technologies de Lille, Co-Directeur de Recherche

Avant-propos

Le travail présenté dans cette thèse a été mené au sein de l'équipe *Scènes Dynamiques* du Laboratoire d'Automatique I^3D ¹ de l'Université des Sciences et Technologies de Lille 1.

Je tiens à remercier Monsieur **Olivier COLOT**, Professeur à l'USTL, pour m'avoir fait l'honneur de présider le jury de cette thèse.

Je tiens à exprimer toute ma gratitude à Monsieur le Professeur **Jack-Gérard POSTAIRE**, mon Co-Directeur de recherche, pour l'intérêt qu'il a apporté à mes travaux, pour la confiance qu'il m'a accordée. Ses conseils, ses constants encouragements et sa grande disponibilité m'ont permis de mener à bien ce travail.

Ma reconnaissance et mes vifs remerciements vont à Monsieur **Yassine RUICHEK**, mon second Co-Directeur de recherche, Maître de Conférences à l'Université de Technologie de Belfort-Montbéliard. Ses conseils, sa rigueur scientifique ainsi que ses idées originales et novatrices m'ont été très précieux pour mener à bien ces années de recherche.

Je suis très honoré par la présence dans ce jury de Monsieur **Abderrafiaa KOUKAM**, Professeur à l'Université de Technologie de Belfort-Montbéliard et de Monsieur **Abdelaziz BENSRAHAIR**, Professeur à l'Université de Rouen. Qu'ils trouvent ici mes sincères remerciements pour avoir accepté d'être les rapporteurs de ce travail.

Monsieur **Jean-Christophe BURIE**, Maître de Conférences à l'université de la Rochelle, a accepté de juger mon travail. Je lui suis très reconnaissant de faire partie des membres du jury.

Je remercie Monsieur **Louahdi KHOUDOUR**, Ingénieur de recherche à l'Institut National de Recherche sur les Transports et leur Sécurité, pour avoir accepté de juger mon mémoire et faire partie du jury de cette thèse.

Je souhaite remercier vivement tous les membres du laboratoire I^3D et particulièrement les membres de l'équipe *Scènes Dynamiques* avec lesquels j'ai passé des moments inoubliables et qui resteront un excellent souvenir.

¹Interaction, Image et Ingénierie de la Décision

J'adresse également mes remerciements à mes amis pour leur disponibilité, leur aide et le soutien constant tout au long des étapes de ce travail.

Les derniers remerciements vont à ma famille pour tout ce qu'elle a fait pour moi et sans laquelle rien de ce qui est entre vos mains aujourd'hui n'aurait été réalisé.

Table des matières

Chapitre 1. Introduction	9
1.1 De la vision humaine à la vision artificielle	9
1.2 La perception du relief	10
1.3 La vision artificielle tridimensionnelle	11
1.4 Organisation du mémoire	12
Chapitre 2. Vision stéréoscopique	14
2.1 Introduction	14
2.2 Stéréovision binoculaire	16
2.2.1 Système de prise de vue	17
2.2.2 Géométrie des capteurs	17
2.2.3 Droites épipolaires	18
2.2.4 Configuration géométrique idéale	19
2.2.5 Paramètres intrinsèques	20
2.2.6 Triangulation géométrique	20
2.2.7 Reconstruction 3D	21
2.3 Stéréovision linéaire	22
2.4 Triangulation géométrique en stéréovision linéaire	24
2.5 Choix de primitives	25
2.6 Extraction des primitives	27
2.6.1 Choix des primitives	27
2.6.2 Extraction des primitives	28
2.6.2.1 Détecteur optimal	28
2.6.2.2 Le détecteur de Deriche	30
2.6.2.2.1 Le détecteur de Canny	31

2.6.2.2.2	Le détecteur optimal de contours de Deriche	31
2.6.2.3	Extraction des points de contour	33
2.6.3	Conclusion	35
2.7	Critères de ressemblance et contraintes	35
2.7.1	Critères de ressemblance	36
2.7.2	Contraintes stéréoscopiques	36
2.8	Mise en correspondance	37
2.8.1	Méthodes corrélatives	37
2.8.2	Méthodes basées sur la programmation dynamique	37
2.8.3	Appariement par prédiction et vérification d'hypothèses	38
2.8.4	Méthodes séquentielles	39
2.8.4.1	Méthode bidirectionnelle	39
2.8.4.2	Méthode hiérarchique	40
2.8.5	Méthodes basées sur les réseaux de neurones	40
2.8.6	Approche par optimisation utilisant le recuit simulé	41
2.8.7	Approche par algorithmes génétiques	42
2.8.8	Algorithme des mouches dynamiques (approche parisienne)	43
2.8.9	Évaluation de la carte dense de disparité par algorithmes génétiques	44
2.9	Conclusion	46
Chapitre 3. Approche globale pour l'appariement d'images stéréoscopiques		48
3.1	Introduction	48
3.2	Algorithmes génétiques	49
3.2.1	Méthodes de sélection	50
3.2.2	Opérateur de croisement	51
3.2.3	Opérateur de mutation	52
3.2.4	Convergence des AGs	53
3.2.5	Conclusion	53
3.3	Codage binaire du problème de mise en correspondance	54
3.4	Contraintes de mise en correspondance	57
3.4.1	Contraintes locales	57
3.4.1.1	Contrainte de position	57

3.4.1.2	Contrainte de signe de la dérivée	57
3.4.2	Contraintes globales	58
3.4.2.1	Contrainte d'unicité	58
3.4.2.2	Contrainte d'ordre	58
3.4.2.3	Contrainte de continuité de la disparité	59
3.4.3	Conclusion	60
3.5	Amélioration du codage	60
3.6	Fonction d'évaluation	61
3.6.1	Fonction d'évaluation représentant la contrainte d'unicité	61
3.6.2	Fonction d'évaluation représentant la contrainte d'ordre	63
3.6.3	Fonction d'évaluation représentant la contrainte de continuité	64
3.6.4	Fonction d'évaluation totale	65
3.7	Mise en œuvre de l'algorithme génétique	66
3.8	Résultat de la mise en correspondance	68
3.8.1	Choix des paramètres	69
3.8.1.1	Choix des coefficients intervenant dans la fonction d'évaluation	69
3.8.1.2	Choix des paramètres génétiques	70
3.8.2	Séquence "Piéton"	71
3.8.2.1	Description et extraction des points de contour	71
3.8.2.2	Résultat de la mise en correspondance	75
3.8.2.3	Conclusion	84
3.8.3	Séquence "Approche d'un véhicule"	84
3.8.3.1	Description et extraction des points de contour	84
3.8.3.2	Résultat de la mise en correspondance	87
3.9	Conclusion	91
Chapitre 4. Nouveau codage génétique pour l'appariement d'images stéréoscopiques		93
4.1	Introduction	93
4.2	Nouveau codage génétique	94
4.3	Exemple	96
4.4	Adaptation de la fonction d'évaluation	99

4.5	Mise en correspondance pour la séquence "Piéton" avec codage entier et fonction d'évaluation adaptée	100
4.5.1	Conclusion	102
4.6	Nouvelle fonction d'évaluation des chromosomes entiers	103
4.7	Résultat obtenu sur la séquence "Piéton" grâce au codage entier et la nouvelle fonction d'évaluation	104
4.8	Résultat obtenu sur la séquence "Approche d'un véhicule"	106
4.9	Réglage des paramètres $Taille_{pop}$ et N_{gen}	107
4.10	Autres résultats expérimentaux	112
4.10.1	Traitement du couple d'images "Voie"	113
4.10.1.1	Description et extraction des points de contour	113
4.10.1.2	Résultat de la mise en correspondance	114
4.10.2	Traitement du couple d'images "Couloir"	115
4.10.2.1	Description et extraction des points de contour	115
4.10.2.2	Résultat de la mise en correspondance	116
4.10.3	Traitement du couple d'images "Village"	117
4.10.3.1	Description et extraction des points de contour	117
4.10.3.2	Résultat de la mise en correspondance	119
4.10.4	Traitement du couple d'images "Rails"	120
4.10.4.1	Description et extractions des points de contour	120
4.10.4.2	Résultat de la mise en correspondance	120
4.11	Conclusion	123
Chapitre 5. Stratégie de recherche hiérarchique		125
5.1	Introduction	125
5.2	Réduire la combinatoire du problème	126
5.3	Analyse du problème	127
5.4	Principe de base de la méthode hiérarchique	128
5.5	Description de la méthode hiérarchique	130
5.5.1	Définitions	130
5.5.2	Algorithme génétique de mise en correspondance et méthode hiérarchique	131
5.6	Implantation récursive de la méthode hiérarchique	132

5.6.1	Principe de base	132
5.6.2	Exploration de l'arbre	134
5.6.3	Prise en compte des résultats de la procédure récursive	135
5.6.4	Implantation récursive	136
5.6.4.1	Sélection des points de contour les plus significatifs	136
5.6.4.2	Initialisation de l'algorithme	138
5.6.4.3	Conditions d'arrêts	139
5.6.4.3.1	Condition d'arrêt 1	139
5.6.4.3.2	Condition d'arrêt 2	139
5.6.4.4	Module d'appariement élémentaire hiérarchique	140
5.7	Application du schéma hiérarchique de mise en correspondance	142
5.7.1	Séquence stéréoscopique "Piéton"	143
5.7.2	Séquence stéréoscopique "Approche d'un véhicule"	146
5.7.3	Conclusion	148
5.8	Conclusion	149
Chapitre 6. Conclusion générale		151
Bibliographie		164

Chapitre 1

Introduction

L'être humain s'est toujours efforcé d'améliorer ses conditions de vie. Pour cela, il n'a cessé de faire progresser la science afin de satisfaire ses intérêts. C'est ainsi que pour remédier aux troubles de la vue, il créa, dès le 11^{me} siècle, de façon rudimentaire la loupe. Puis ses connaissances en optique lui permirent, à partir du 13^{me} siècle, de mettre au point les lunettes. Mais l'Homme, dans sa soif de connaissances, n'est pas arrêté à ses déficiences, il a voulu repousser les limites qui lui imposait son système visuel. Ainsi, le monde de l'infiniment petit s'ouvrit à lui avec, à la fin du 16^{me} siècle, l'invention du microscope par l'opticien hollandais H. Jansen. Galilée, quand à lui, repoussa les limites à l'infini en inventant, en 1609, la lunette astronomique. Les objets les plus lointains étaient alors à "portée d'œil". La science a également permis d'aller jusqu'à recréer artificiellement le processus de la vision.

1.1 De la vision humaine à la vision artificielle

Qu'est-ce que la vision ? Pour répondre à cette question, nous sommes tentés de dire : la vision c'est voir ce qui nous entoure. En effet, l'Homme est doté d'un organe, l'œil, chargé de percevoir les phénomènes lumineux afin de recréer une image du monde extérieur. Cette image se forme sur une membrane, la rétine, située dans le fond de l'œil et constituée de centaines de milliers de cellules nerveuses. L'image projetée sur celle-ci n'est donc ni plus ni moins qu'un nombre impressionnant de points. Mais cette définition de la vision est incomplète car voir c'est également interpréter et comprendre. Comment, à partir de cette multitude de points, sommes-nous capables d'identifier une chaise ou une automobile ?

Si la formation d'une image est une condition nécessaire à la vision, elle n'est pas suffisante. Il faut encore que l'organe récepteur soit solidaire d'un système nerveux suffisamment

développé pour traiter l'information visuelle. Les cellules nerveuses composant la rétine ne se contentent pas de recevoir les stimuli lumineux. Celles-ci, directement reliées au cerveau, lui transmettent par l'intermédiaire des fibres du nerf optique les radiations lumineuses sous la forme d'un influx nerveux. C'est alors au niveau du cortex cérébral que sont interprétées les sensations reçues pour nous renseigner sur le monde extérieur.

L'œil et le cerveau sont ainsi capables de percevoir les formes, les couleurs et les mouvements des objets qui nous entourent. Ils sont aussi capables grâce au système visuel binoculaire, de déterminer la géométrie spatiale de l'environnement. L'Homme peut ainsi se déplacer, attraper un objet, ou contourner un obstacle en évaluant la taille et la distance des éléments qui l'entourent.

Les images vues par l'œil droit et par l'œil gauche ne sont pas identiques. Pour reconstruire la troisième dimension, le cerveau utilise le décalage existant entre les deux images rétiniennes, appelé disparité stéréoscopique. Il s'agit alors pour celui-ci d'établir des correspondances entre les deux images d'un même objet.

Cependant, on comprend encore assez mal comment ces correspondances sont établies par le cerveau. D'ailleurs ce sujet de recherche constitue un domaine très actif en vision artificielle visant à créer des systèmes intelligents dotés d'un système visuel semblable à celui de l'Homme.

1.2 La perception du relief

La perception du relief est une fonction fondamentale qui permet de structurer le monde et l'espace environnant, d'évaluer les grandeurs et les distances. Mais cette perception est personnelle, rien ne nous permet de la partager. De là est né le besoin de trouver le moyen de restituer fidèlement la réalité de l'espace.

Autrefois, l'Homme ne disposait, pour représenter le monde qui l'entoure, que d'un support à deux dimensions. Mais restituer sur un seul plan ce que l'œil perçoit en trois dimensions est un exercice difficile.

La première tentative pour obtenir une image en relief fut réalisée par Charles Wheatstone en 1838. Celui-ci eut l'idée d'utiliser deux images et conçut le premier stéréoscope. Cet appareil permet d'observer simultanément deux images d'un même objet, prises sous deux angles différents. Il restitue ainsi la profondeur et le relief.

En 1891, Louis Ducos de Hauron développa un nouveau système, les anaglyphes, qui consiste

à superposer deux images différentes de la même vue après avoir teintées l'une en bleu-violet, l'autre en rouge, puis à les regarder à l'aide de lunettes ayant respectivement des verres de mêmes couleurs. Chaque œil n'aperçoit que l'image qui lui convient et le cerveau assure la fusion des deux photographies pour donner l'illusion d'une seule image en relief.

Si le stéréoscope et les anaglyphes sont toujours employés, l'avenir appartient aux techniques holographiques. Ces techniques utilisent les interférences produites par la superposition de deux faisceaux laser pour reproduire une photographie en relief et permettent ainsi de représenter fidèlement les objets en trois dimensions.

De nos jours, on ne se contente plus de vouloir restituer le relief. Les chercheurs s'intéressent également au moyen de percevoir le relief grâce à la vision artificielle.

1.3 La vision artificielle tridimensionnelle

Depuis quelques années, la vision par ordinateur a fait des progrès considérables. Le plus remarquable d'entre eux est la possibilité d'obtenir des informations tridimensionnelles à partir d'images vidéo.

En effet, en disposant de plusieurs images prises sous des angles différents, il devient possible de calculer la position spatiale des points observés dans au moins deux images. On appelle **stéréoscopie** ou **stéréovision** le processus qui permet de combiner entre elles plusieurs images d'une même scène pour en extraire des informations géométriques tridimensionnelles.

La grande difficulté de la vision artificielle réside dans l'interprétation des images. Dans le cas de la stéréovision, celle-ci est encore plus accrue. En effet, la perception du relief n'est possible que si l'on réalise, comme le fait le cerveau, l'opération de fusionnement des deux images. En vision par ordinateur, l'étape de fusionnement porte le nom de **mise en correspondance** et fait l'objet de nombreuses recherches. Cette thèse en fait partie.

En effet, le laboratoire I^3D est l'un des acteurs qui ont participé à différents projets de recherches dans le but d'améliorer certains aspects du trafic routier (PROMETHEUS, STATUE, etc.). À titre d'exemple, dans le cadre du projet PROMETHEUS¹ lancé par les ministres européens des affaires étrangères, de l'économie et de la recherche, le laboratoire a conçu, mis en œuvre et validé un capteur stéréoscopique constitué de deux caméras linéaires afin de détecter les obstacles sur route [Bru94]. Ensuite, Jean Christophe Burie et Yassine Ruichek ont

¹PROgramme for a European Traffic with Highest Efficiency and Unprecedented Safety.

développé deux approches pour résoudre le problème de l'appariement des primitives extraites d'images stéréoscopiques. La première est une approche locale qui recherche les appariements couple par couple en explorant séquentiellement les couples d'images stéréoscopiques [Bur95] alors que la deuxième est basée sur l'utilisation des réseaux de neurons de Hopfield [Rui97].

Les techniques de mise en correspondance séquentielles ou multi-résolutions ayant montré leurs limites [BBP95], nous proposons une technique d'optimisation globale basée sur les concepts du calcul génétique. Cette approche fait appel à une fonction d'évaluation qui intègre toutes les contraintes qui permettent de lever les ambiguïtés d'appariement. Grâce à un codage spécifique permettant une représentation des solutions sous forme de chromosomes, le processus de mise en correspondance consiste à optimiser la fonction d'évaluation en utilisant des opérations de sélection et de reproduction [Gol89]. On recherche ainsi, par une exploration de l'espace des solutions, l'optimum de la fonction d'évaluation qui correspond aux meilleurs appariements entre les primitives des images stéréoscopiques. Notre approche est similaire à d'autres techniques utilisant les réseaux de neurons de Hopfield [RP99]. L'avantage de l'approche par calcul génétique réside dans sa nature de recherche globale de l'optimum, contrairement aux réseaux de Hopfield qui ne mettent en œuvre qu'un processus d'optimisation locale.

1.4 Organisation du mémoire

Dans ce mémoire nous nous intéressons particulièrement au problème de l'appariement des indices visuels, appelés également **primitives**, extraits de couples d'images stéréoscopiques acquises par des stéréoscopes placés selon une configuration particulière.

Dans le chapitre 2, nous rappelons les principes fondamentaux de la vision stéréoscopique binoculaire. Nous commençons, tout d'abord, par présenter les différentes configurations géométriques pour constituer un stéréoscope. Nous présentons tout particulièrement les configurations des stéréoscopes développés au laboratoire d'Automatique I^3D pour la détection d'obstacles. Nous décrivons ensuite la méthode d'extraction des primitives utilisées lors de la phase de mise en correspondance. À la fin de ce chapitre, nous exposons les principes de base de quelques méthodes de mise en correspondance en vision stéréoscopique.

Dans le chapitre 3 et avant de présenter notre approche globale de mise en correspondance basée sur l'utilisation des algorithmes génétiques [Gol89], nous rappelons succinctement les différentes phases de traitement de ces derniers : sélection des individus, croisement et mu-

tation. Suite à ce rappel, nous décrivons le codage élaboré, les contraintes stéréoscopiques à partir desquelles la fonction d'évaluation des chromosomes est définie. Ensuite, nous validons notre approche génétique sur deux séquences d'images stéréoscopiques réelles acquises grâce au stéréoscope linéaire développé au laboratoire I^3D .

Nous proposons, dans le chapitre 4, un nouveau codage génétique ainsi qu'une nouvelle fonction d'évaluation pour résoudre plus efficacement le problème de mise en correspondance. Un AG, doté de ce nouveau codage et de cette nouvelle fonction d'évaluation, est validé d'une part sur des images stéréoscopiques acquises grâce au stéréoscope linéaire et, d'autre part, sur des images stéréoscopiques de synthèse ainsi que sur des images réelles acquises par un stéréoscope constitué, cette fois-ci, de deux caméras matricielles. Ce dernier a été développé dans le cadre du projet STATUE² dont un des principaux objectifs est de résoudre le problème de la détection d'obstacles devant un véhicule guidé, tel qu'une rame de métro.

Afin d'envisager une exploitation effective de notre procédure génétique et notamment pour des applications nécessitant des réponses en "temps-réel", nous présentons, dans le chapitre 5, une méthode hiérarchique permettant de décomposer le problème en plusieurs sous-problèmes de combinatoires moins élevées. Cette méthode est testée sur les mêmes séquences d'images acquises avec le stéréoscope linéaire et les résultats de mise en correspondance obtenus sont comparés avec ceux obtenus par l'approche génétique globale.

En conclusion de ce mémoire, nous présentons les améliorations susceptibles d'être apportées à notre approche de mise en correspondance, ainsi que quelques perspectives futures.

²Système Télématique destiné à l'Accroissement de la sûreté des Transports guidés Urbains et à l'Émergence de nouvelles aides à l'exploitation.

Chapitre 2

Vision stéréoscopique

2.1 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 en fonction de notre environnement. Nous sommes ainsi capables de percevoir les formes, les couleurs et les mouvements des objets. Mais nous possédons aussi la capacité de structurer le monde extérieur, d'évaluer les grandeurs et les distances grâce à notre système visuel binoculaire.

Avec l'apparition des machines de calcul de plus en plus sophistiquées, un certain nombre de scientifiques se sont attaqués au problème de la vision d'un point de vue quantitatif. Il ne s'agit pas de fournir une explication de la vision biologique, mais de créer des modèles qui peuvent reproduire le même type de performances visuelles. Grâce aux efforts des chercheurs en robotique et en intelligence artificielle, il est possible aujourd'hui de concevoir des systèmes de vision capables d'effectuer certaines tâches sous des conditions bien contrôlées, comme la reconnaissance de certains objets sur une chaîne de production, mais on est encore loin de fournir un modèle qui parviendra à rivaliser avec la vision humaine.

Un des objectifs de la vision artificielle est de percevoir le relief d'un environnement à partir d'une ou plusieurs images. La vision stéréoscopique constitue l'une des méthodes les plus utilisées dans ce domaine. Il s'agit de retrouver la géométrie spatiale d'une scène à partir de plusieurs images prises sous des angles différents. Le système stéréoscopique le plus courant et le plus simple n'utilise que deux images de la scène observée. D'autres systèmes stéréoscopiques utilisent plus de deux caméras [KO94, Cox94, Fua95]. On peut citer par exemple, la stéréovision trinoculaire qui est basée sur l'utilisation de 3 vues de la scène [YKK86]. Par rapport à la stéréovision binoculaire, la stéréovision trinoculaire permet de lever certaines ambiguïtés en

cas d'occlusions. En revanche, le champs commun aux 3 caméras se trouve réduit par rapport à celui formé par deux caméras.

Les domaines d'application de la stéréovision sont nombreux. Nous pouvons citer par exemple :

- La reconstruction de terrains à partir d'images satellitaires ;
- La modélisation de scènes d'intérieur ou d'extérieur ;
- L'analyse d'images médicales et biomédicales ;
- La perception dynamique pour robots mobiles ;
- Le contrôle de production en $3D$;
- La détection d'obstacles.

Comme chez l'homme, la perception du relief 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. Le problème essentiel de la vision binoculaire est donc de trouver dans les deux vues les images des mêmes entités physiques de la scène observée. Il faut donc apparier les deux images. Sans contraintes particulières, une mise en correspondance point à point des images n'est pas une solution idéale. En effet, avec des images de taille courante de 512×512 pixels, on dénombre 512^4 , soit 68 milliards couples de points à tester. De plus, tout le contenu d'une image n'est pas forcément porteur d'informations pertinentes. On procède donc à l'extraction de primitives telles que points remarquables, contours, régions, etc. Ces primitives-image $2D$ doivent correspondre à des primitives-objet $3D$ dans la scène. Ainsi, apparier les images revient à apparier les primitives-image. Constituant une description de la scène perçue, les primitives-objet peuvent être reconstruites en utilisant des calculs de triangulation géométrique sur les paires de primitives-image homologues. Ces calculs nécessitent la définition de coordonnées dans les images et dans la scène, et donc de connaître la géométrie du capteur stéréoscopique.

Les méthodes de mise en correspondance peuvent être grossièrement classées en deux catégories :

- *Les méthodes denses*¹ : Cette catégorie de méthodes est basée sur la mise en correspondance des pixels des deux images stéréoscopiques et fournissent des cartes denses de disparité². La recherche des appariements est basée sur un critère de ressemblance photométrique. Cette recherche consiste à considérer une fenêtre autour du pixel que l'on veut

¹Area-based methods ou intensity-based methods

²Dense disparity maps

apparié, puis à chercher dans l'autre image le pixel dont la fenêtre, de même taille que la précédente, est la mieux corrélée [CM99, KO94, LOY73, ZK00].

- *Les méthodes non-denses*³ : Les méthodes de cette catégorie sont basées sur la mise en correspondance de primitives extraites des deux images stéréoscopiques et fournissent des cartes non denses de disparité⁴. Contrairement au pixel qui est caractérisé uniquement par un niveau de gris, les primitives mettent en évidence d'autres caractéristiques qui permettent de lever certaines ambiguïtés de mise en correspondance. Ces méthodes sont généralement utilisées dans des applications ne nécessitant pas une reconstruction 3D complète (pixel par pixel) de la scène observée, telles que la détection d'obstacles [Gri81, Gri85, HSC⁺89, Bur95, Rui97].

Dans le cadre de cette thèse nous nous intéressons à la deuxième catégorie de méthodes stéréoscopiques en utilisant des primitives de type point de contour. La suite de ce chapitre donne une présentation du principe de la stéréovision depuis le système de prise de vue jusqu'à la reconstruction 3D.

2.2 Stéréovision binoculaire

La stéréovision binoculaire est un processus qui, à partir de deux images représentant la même scène sous des angles différents, permet de retrouver la structure spatiale de celle-ci. Ce processus peut être décomposé en quatre étapes :

- Choix d'un modèle géométrique pour les capteurs ;
- Extraction des primitives de chaque image ;
- Appariement des primitives ;
- Triangulation géométrique et reconstruction tridimensionnelle de la scène.

Dans ce processus, la tâche la plus difficile et la plus délicate est celle concernant la mise en correspondance des primitives extraites de deux images stéréoscopiques. En effet, cette étape est de nature combinatoire car elle dépend du nombre de primitives candidates à l'appariement. Il est donc fondamental de réduire la combinatoire, d'une part pour diminuer les risques d'erreurs d'appariement, mais également pour limiter le temps des traitements. Cependant, la seule façon de réduire cette combinatoire est de mettre en œuvre des contraintes, notamment

³Feature-based methods

⁴Sparse maps

la contrainte épipolaire qui résulte de la géométrie du capteur stéréoscopique. L'application de cette contrainte toute seule permet, comme on va le voir dans la suite, de réduire considérablement le nombre d'appariements possibles.

Après la phase de mise en correspondance, les couples de primitives-image appariées permettent, par triangulation, de calculer les positions des primitives-objet correspondantes dans la scène réelle. Comme on va le voir plus loin, les calculs de triangulation sont liés à la géométrie du système de prise de vue, plus particulièrement à l'entre-axe, c'est-à-dire la distance séparant les centres optiques des deux caméras. La précision de la localisation est accrue si les points de vue sont nettement différents, c'est-à-dire lorsque l'entre-axe est important. 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 occlusions sont également plus probables. Par conséquent, la phase 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.

2.2.1 Système de prise de vue

2.2.2 Géométrie des capteurs

La géométrie des capteurs est un élément important du dispositif de prise de vue stéréoscopique. Il s'agit soit de déterminer les paramètres caractérisant une disposition quelconque des caméras (cf. figure 2.1), soit au contraire d'imposer une configuration particulière aux capteurs en imposant les paramètres voulus (cf. figure 2.2). Ces paramètres sont ensuite utilisés lors de la recherche des couples à appairier.

Considérons le schéma de la figure 2.1 qui représente deux caméras. Pour qu'elles observent la même scène, les deux caméras doivent avoir une configuration telle que leur axes optiques convergent vers la scène observée. A chaque caméra est associé un repère. Soit P un point de la scène et soient p_g et p_d ses deux projections dans les images gauche et droite (p_g et p_d sont dits des points homologues). Nous pouvons donc écrire l'équation de la droite passant par le centre focal O_g de la caméra gauche et le point p_g . De même, on peut écrire l'équation de la droite passant par le centre focal O_d de la caméra droite et le point 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 pouvoir 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. Ici apparaît la nécessité de connaître les relations géométriques permettant de passer du repère lié à la caméra gauche au repère lié à la caméra droite, ainsi que celles permettant de passer des repères-caméras au repères associé à la scène. La détermination de ces relations s'effectue lors de la phase de calibrage du système de prise de vue stéréoscopique.

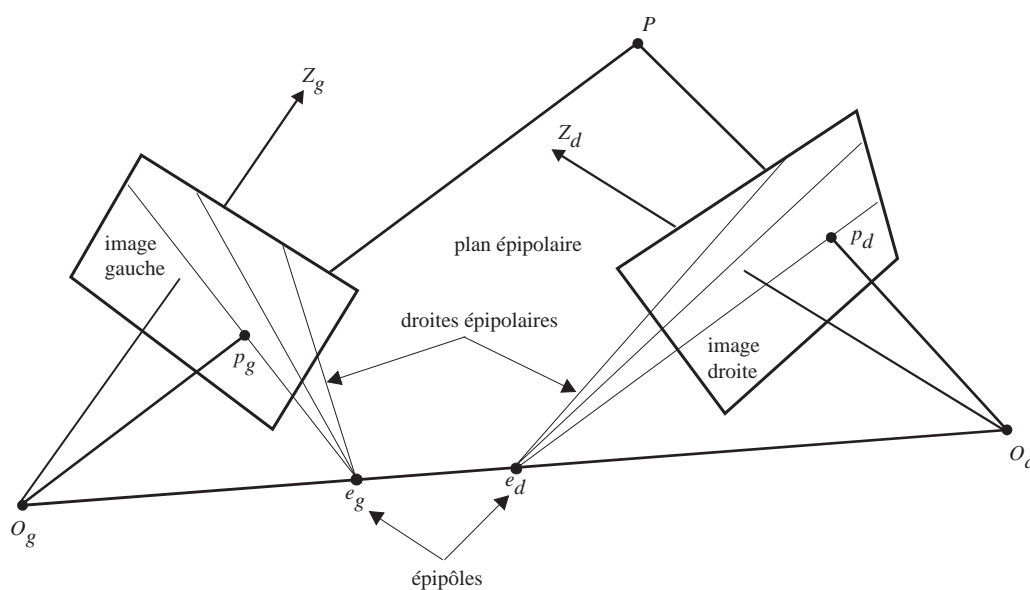


FIG. 2.1 : Configuration générale d'un capteur stéréoscopique.

2.2.3 Droites épipolaires

Lorsque les paramètres extrinsèques du capteur stéréoscopique qui définissent la position et l'orientation de chaque caméra par rapport à la scène sont connus, ceux-ci sont utilisés dans la phase de mise en correspondance. Ils font appel à la notion de droite épipolaire qui est le lieu des points d'une image pouvant correspondre à un même point p de l'autre image (cf. figure 2.1). 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 y a une droite épipolaire dans l'image gauche. De plus, deux points qui se trouvent sur une même droite épipolaire d'une image ont leurs correspondants, s'ils existent, sur une même 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, et leur point commun s'appelle l'épipôle (cf. figure 2.1). En fait, l'épipôle droit e_d , par exemple, n'est que la projection du centre focal O_g de la caméra gauche

dans l'image droite. De la même manière on peut définir l'épipôle gauche e_g .

2.2.4 Configuration géométrique idéale

Dans le cas d'une configuration quelconque du capteur stéréoscopique (cf. figure 2.1), 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 particulière intéressante : c'est lorsque les deux caméras sont disposées de telle façon que leurs axes optiques sont parallèles et lorsque la droite $O_g O_d$ reliant les centres optiques est confondue avec les axes horizontaux des deux caméras (cf. figure 2.2). Autrement dit, lorsque les droites épipolaires sont parallèles et les droites épipolaires conjuguées sont confondues. Dans ce cas, les deux épipôles se trouvent à l'infini.

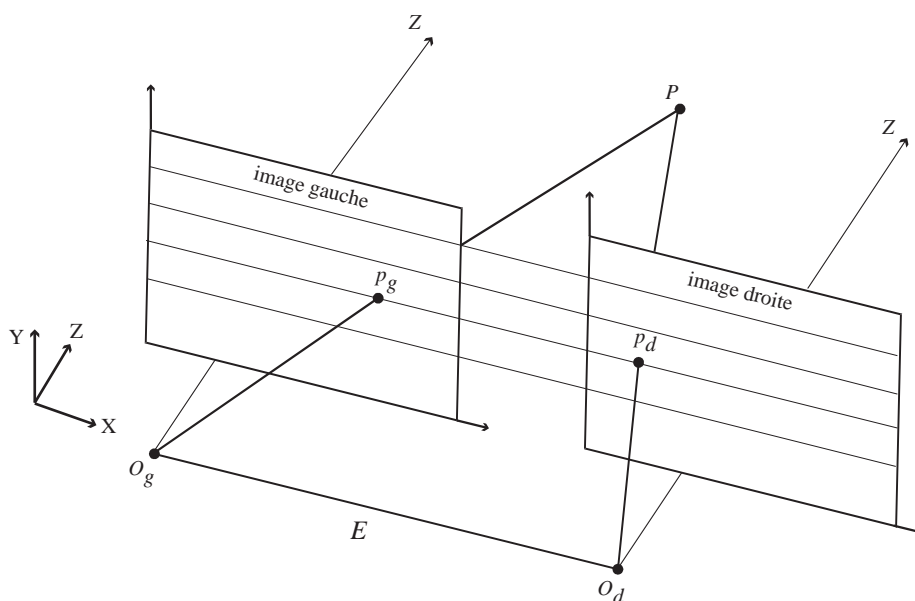


FIG. 2.2 : Configuration particulière du système de prise de vue.

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 particulière de la figure 2.2, de nombreux auteurs [HM93, Aya89] préfèrent partir d'une configuration quelconque (cf. figure 2.1) et se ramener de façon analytique à la configuration particulière (cf. figure 2.2) en effectuant une rectification épipolaire. Celle-ci consiste à appliquer aux images une transforma-

tion 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 [Las96].

2.2.5 Paramètres intrinsèques

Outre les paramètres extrinsèques qui définissent la géométrie du capteur stéréoscopique, en particulier la position et l'orientation de chaque caméra par rapport à la scène, d'autres paramètres doivent être choisis judicieusement afin d'optimiser, en fonction de l'application, le système de vision. Il s'agit des paramètres correspondant aux caractéristiques propres des caméras :

- Taille et résolution des capteurs ;
- Distance entre les centres optiques ;
- Distance focale des objectifs.

Les deux derniers paramètres interviennent lors des calculs de triangulation géométrique. Ils permettent aussi de déterminer :

- La taille de la scène observable à une distance donnée ;
- La distance minimale de prise de vue permettant d'appliquer le principe de stéréovision.

2.2.6 Triangulation géométrique

Les calculs de triangulation géométrique permettent, en fonction des positions dans les images des différentes primitives appariées, de déterminer la position des primitives-objet correspondantes dans la scène tridimensionnelle. Cette procédure utilise un modèle géométrique du système de prise de vue, décrivant la formation des images sur les capteurs. Considérons la configuration particulière du stéréoscope présentée au paragraphe 2.2.4 (cf. figure 2.3). Soient $p_g(x_g, y_g)$ et $p_d(x_d, y_d)$ les coordonnées des images du point P sur les capteurs gauche et droit, dans les repères relatifs aux caméras.

Comme, dans ce cas, les droites épipolaires sont confondues ($y_g = y_d$), la position de p_g dans l'image gauche peut se mesurer par rapport à celle de p_d avec un seul paramètre, la disparité, qui est par définition la différence des abscisses des deux points -image. La disparité entre les deux points homologues p_g et p_d est notée : $\delta = x_g - x_d$. Ainsi, la position du point

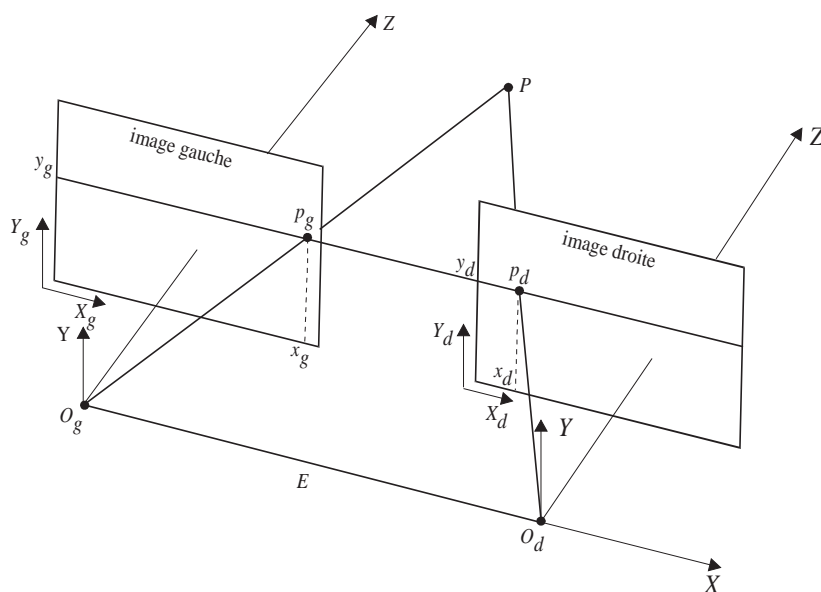


FIG. 2.3 : Localisation tridimensionnelle.

P dans le repère absolu (O_g, X, Y, Z) peut être déterminée par les équations suivantes :

$$\begin{aligned}x_p &= \frac{x_g \cdot E}{\delta} \\y_p &= \frac{y_g \cdot E}{\delta} \\z_p &= \frac{f \cdot E}{\delta}\end{aligned}$$

où E est l'entre-axe, c'est-à-dire la distance entre les centres optiques O_g et O_d .

f est la distance focale des objectifs, supposés identiques pour les deux caméras.

2.2.7 Reconstruction 3D

Pour reconstruire l'image 3D de la scène observée, on utilise les positions tridimensionnelles des primitives-objet calculées par triangulation. Il faut noter que la reconstruction 3D dépend du type de primitive choisi, qui détermine 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 la combinatoire. En revanche, lorsque les primitives ne sont pas nombreuses, la combinatoire est réduite mais la reconstruction est moins riche.

2.3 Stéréovision linéaire

Comme pour la stéréovision binoculaire classique, la stéréovision linéaire désigne l'ensemble de tous les procédés qui permettent de retrouver la structure tridimensionnelle d'une scène à partir de deux images vidéo prises sous des angles de vue différents. La différence fondamentale entre les deux systèmes stéréoscopiques réside dans le type de caméra utilisée pour acquérir les images. En effet, dans le cas de la stéréovision linéaire, le stéréoscope est composé de deux caméras linéaires plutôt que matricielles. La meilleure façon d'imaginer une caméra linéaire est de penser à une caméra matricielle pour laquelle une seule ligne serait active. Physiquement, une caméra linéaire comporte une surface sensible monodimensionnelle constituée d'une barrette CCD⁵. Munie de son optique, elle fournit une image ligne qui est une vue en coupe de la scène observée.

Ces capteurs linéaires présentent certaines caractéristiques qui les rendent particulièrement intéressants dans de nombreuses applications :

- La grande résolution horizontale : leur structure très simple permet d'intégrer un grand nombre de photosites en augmentant la longueur de la barrette. La résolution horizontale typique est de 2048 points, contre environ 750 points pour les capteurs matriciels courants.
- La compatibilité avec les optiques photographiques : la longueur du capteur est généralement comprise entre 20 et 28mm, ce qui permet d'utiliser des objectifs photographiques dont les caractéristiques de résolution, de distorsion et de vignétage sont très supérieures à celles des objectifs vidéo. La grande résolution des objectifs photographiques est indispensable pour tirer parti de celle des capteurs linéaires. Leur faible distorsion peut généralement être négligée, ce qui évite des calculs de correction d'images. Enfin, leur faible vignétage évite des problèmes lors de l'appariement des images en stéréovision.
- La souplesse des paramètres de prise de vue : l'opérateur peut ajuster, selon les besoins et les conditions de prise de vue, la fréquence d'acquisition et le temps d'exposition. Il peut synchroniser chaque prise de vue sur un événement extérieur, même à intervalles irréguliers. Il est même possible de définir la fréquence de transfert des pixels en fonction des performances de la chaîne d'acquisition et de traitement des images.

Il faut noter que les capteurs linéaires posent cependant certains problèmes qui sont liés à

⁵Charge Coupled Device

leur structure et au type de signal qu'ils fournissent. Il s'agit de :

- La difficulté du calibrage : lors de l'utilisation de caméras linéaires en stéréovision, il est nécessaire d'avoir un positionnement précis et rigide des deux caméras afin qu'elles observent une zone de vision commune.
- La pauvreté relative du signal monodimensionnel : à cause de la perte de l'information verticale, il est difficile d'interpréter une image-ligne. La dimension manquante peut-être cependant reconstruite, dans certains cas, grâce à un balayage mécanique de la scène.

Dans le but de détecter les obstacles sur route, le laboratoire *I3D* a conçu un stéréoscope composé de deux caméras linéaires [Bru94]. Ce stéréoscope n'est pas différent dans son principe de tous les systèmes de prise de vue stéréoscopique à deux caméras. Cependant, afin de rendre possible la reconstruction tridimensionnelle, la structure géométrique des caméras linéaires, qui ne restituent pas l'information verticale, impose une configuration particulière pour que les deux caméras du stéréoscope visent la même scène (*cf.* figure 2.4). Pour cela, il faut d'une part, que les deux capteurs soient dans le prolongement l'un de l'autre. D'autre part, il faut que la ligne joignant les deux capteurs soit parallèle à celle joignant les centres optiques des deux objectifs. La partie commune aux plans de visée des caméras définit ce qu'on appelle le secteur de stéréovision. Tout objet qui entre dans ce secteur est vu simultanément par les caméras gauche et droite. Il peut être ainsi être détecté et localisé par triangulation (*cf.* § 2.2.6).

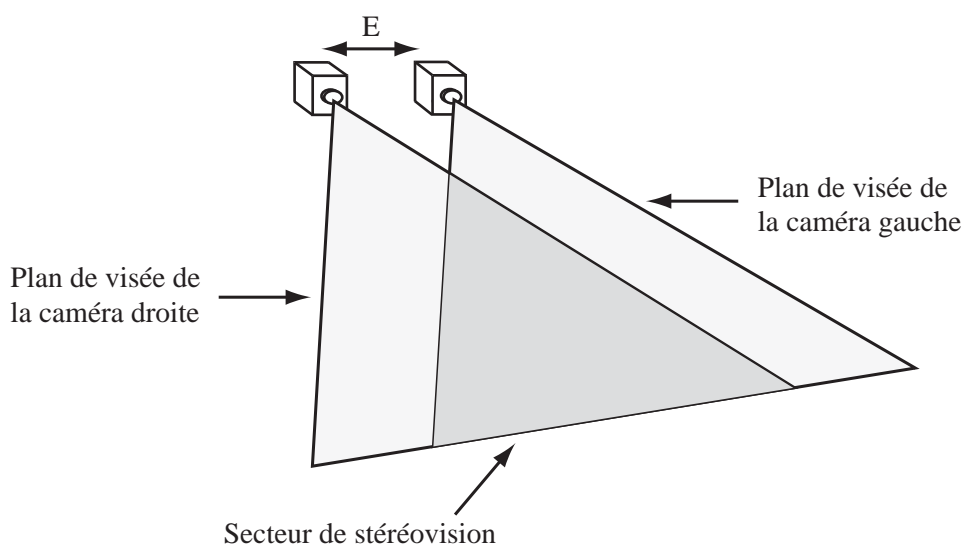


FIG. 2.4 : Configuration du stéréoscope linéaire (les deux plans de visée sont confondus).

2.4 Triangulation géométrique en stéréovision linéaire

Nous rappelons que la triangulation géométrique permet de déterminer les coordonnées tridimensionnelles d'un objet à partir des projections de son image sur les capteurs droit et gauche. Dans le cas de la stéréovision linéaire, le champ de visée est un plan et par conséquent la triangulation géométrique permet de reconstruire un point P uniquement par sa profondeur z et par sa position horizontale x (cf. figure 2.5).

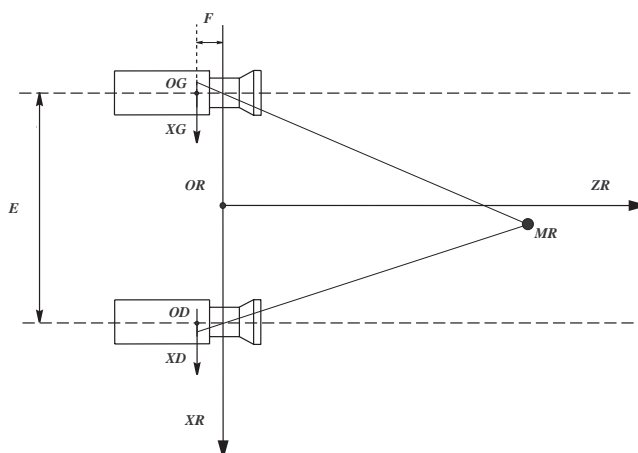


FIG. 2.5 : Triangulation géométrique dans le plan $(O, (X, Z))$

Les positions des points-images x_g et x_d se déduisent de la position du point de la scène $P = (x, z)$ par les relations :

$$x_g = \frac{f(x + \frac{E}{2})}{z} \quad \text{et} \quad x_d = \frac{f(x - \frac{E}{2})}{z} \quad (2.1)$$

On introduit la notion de disparité qui correspond à la différence de position δ entre les points-images :

$$\delta = x_g - x_d \quad (2.2)$$

En remplaçant dans l'équation (2.2) les valeurs de x_g et x_d obtenues en 2.1, on obtient :

$$\delta = \frac{E \cdot f}{z}, \quad \forall x \quad (2.3)$$

Quelle que soit l'abscisse du point P , on peut donc calculer sa position z par :

$$z = \frac{E \cdot f}{\delta} \quad (2.4)$$

En substituant la valeur de z dans l'une ou l'autre des équations (2.1), on obtient la valeur de x par :

$$x = \frac{Zx_g}{f} - \frac{E}{2} \quad \text{ou} \quad x = \frac{Zx_d}{f} + \frac{E}{2} \quad (2.5)$$

En ce qui concerne la coordonnée selon l'axe verticale Y , celle-ci est perdue à cause de la nature monodimensionnelle des caméras. Néanmoins, nous pouvons déterminer la coordonnée verticale d'un point selon l'axe Y dans un cas particulier à savoir, lorsque les objets observés sont placés sur un plan (cf. figure 2.6). Connaissant la hauteur H des caméras, la distance de ce point z par rapport à ces dernières ainsi que la distance maximale z_{max} visée par celles-ci, nous pouvons en déduire la hauteur h de ce point par l'équation suivante :

$$h = \frac{H(z_{max} - z)}{z_{max}} \quad (2.6)$$

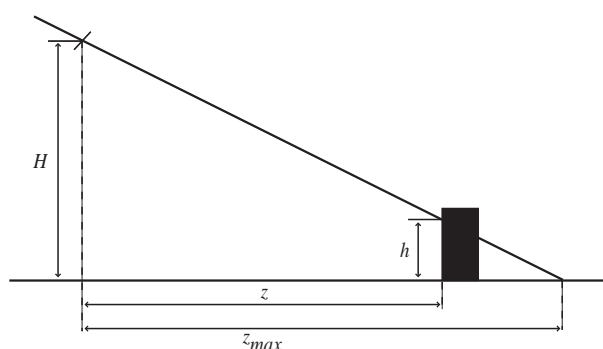


FIG. 2.6 : Cas particulier permettant de déterminer la coordonnée verticale d'un point vu par de caméras linéaires

2.5 Choix de primitives

Nous rappelons que le processus de stéréovision peut être décomposé en quatre étapes :

- Choix d'un modèle géométrique pour les capteurs ;
- Extraction des primitives de chaque image ;
- Appariement des primitives ;
- Triangulation géométrique et reconstruction tridimensionnelle de la scène.

Le choix des primitives est une phase importante dans un algorithme de stéréovision. La nature de ces primitives et les différents attributs qu'on peut leur attribuer jouent un grand rôle

lors de la phase d'appariement mais également dans la reconstruction de la scène. Il est donc nécessaire de choisir les primitives les mieux adaptées au type d'images traitées.

Nous venons de voir que quelque soit le type de stéréoscope utilisé (linéaire ou matriciel), celui-ci est placé dans une configuration particulière afin de limiter l'espace de recherche des points correspondants aux couples de lignes épipolaires. Cette configuration a été choisie afin de réduire le temps nécessaire pour accomplir la phase de mise en correspondance.

Étant donné que les mises en correspondance sont effectuées uniquement sur les lignes épipolaires conjuguées, le choix des primitives se trouve rapidement restreint en raison de la nature monodimensionnelle des images-lignes issues du stéréoscope linéaire ou des lignes épipolaires extraites des images issues du stéréoscope matriciel. En effet, parmi les primitives habituellement utilisées en stéréovision telles que les primitives de type point, contour [SH88, HM93, LL94b, WL95, Xie95], région [dBKS90, PM97] et contour actif [KWT87b, KWT87a, Vie88, Cab92, Sle95], seuls les points de contours non horizontaux sont utilisables. Ces points proviennent de l'intersection des contours de l'image bidimensionnelle avec les droites épipolaires (*cf.* figure 2.7).

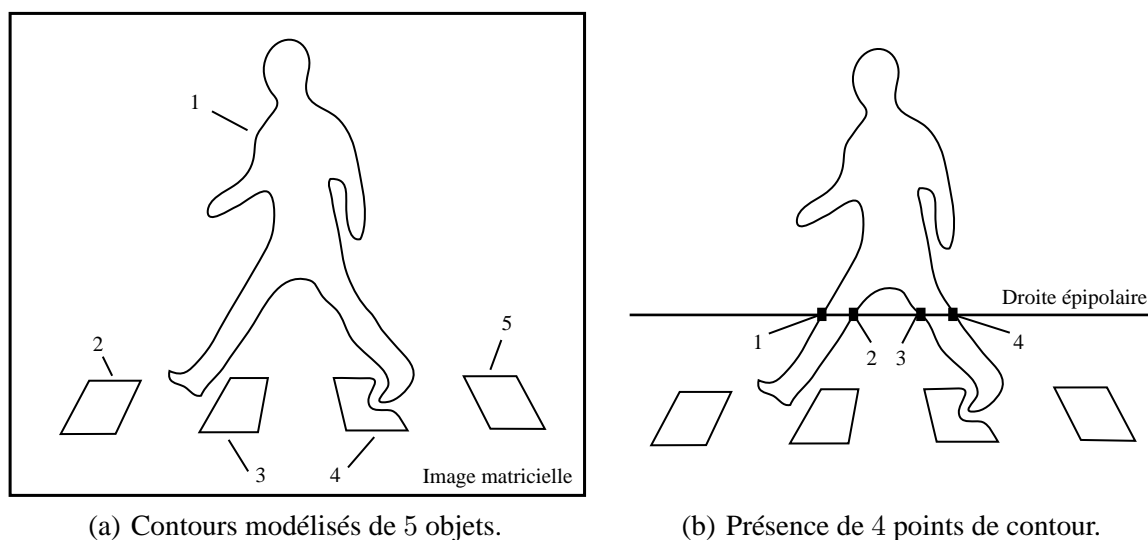


FIG. 2.7 : Comparaison des primitives de type contour.

Dans le paragraphe suivant, nous allons décrire le processus d'extraction de ces primitives de type points de contour, qui constitue la première tâche d'un algorithme de stéréovision.

2.6 Extraction des primitives

Les primitives sont des indices visuels caractérisant des entités physiques présentes dans l'image. L'utilisation de primitives permet de ne retenir que l'information utile pour l'application choisie. A chaque primitive est associé un ensemble d'attributs la caractérisant. Il est nécessaire que ces attributs soient suffisamment discriminants afin d'éviter au maximum les ambiguïtés lors de la phase de mise en correspondance. Par exemple, pour une primitive de type segment, deux attributs couramment utilisés sont la longueur et l'orientation. Horaud [HM95] définit la primitive idéale telle que :

- ses propriétés intrinsèques permettent une mesure de ressemblance fortement discriminante ;
- elle permet la mise en œuvre des contraintes stéréoscopiques de manière efficace ;
- la reconstruction $3D$ est possible.

2.6.1 Choix des primitives

Le choix des primitives mérite une attention particulière. En effet, étant donné la nature combinatoire de la phase de mise en correspondance, si celles-ci sont trop nombreuses, le temps de calcul aura tendance à être important. Il est donc avantageux de réduire le nombre de primitives tout en s'assurant qu'elles apportent une information suffisamment riche pour permettre une reconstruction $3D$ représentative de la scène.

Ce choix peut également être imposé par le type de caméras utilisées. A partir de caméras matricielles, il est d'usage d'utiliser des primitives de type point de Moravec [Mor77], contour, point à forte courbure ou région, dont certains attributs contiennent une information bidimensionnelle [LL90]. En revanche, avec des caméras linéaires, on est contraint d'utiliser des primitives ponctuelles dont les attributs ne représentent que des informations monodimensionnelles.

Dans le cadre de cette thèse, nous avons choisi de travailler sur des couples de lignes épipolaires grâce à l'utilisation de stéréoscopes de configuration particulière (*cf.* figure 2.2). Ainsi, les primitives que nous utilisons ne peuvent être que de type points de contour. Elles sont d'ailleurs couramment utilisées car elles correspondent aux bords des objets ou aux variations spatiales des propriétés géométriques et physiques de leurs surfaces (orientation, matière, texture, niveau de luminance, etc.). Ces éléments significatifs peuvent être facilement extraits des images

(cf. § 2.6.2.2). Comme ils constituent les entrées du processus de mise en correspondance, il est important de s'assurer de leur pertinence (cf. § 2.6.2.3).

2.6.2 Extraction des primitives

L'extraction des primitives se déroule en deux étapes. Les primitives sont d'abord mises en évidence par le biais d'un détecteur de points de contour, puis les plus pertinentes sont retenues par une procédure de filtrage (cf. § 2.6.2.3).

Les points de contour sont caractérisés par des variations locales significatives du niveau de gris $f_{NG}(y)$ de la ligne considérée. La détection des points de contour fait donc appel à des opérateurs différentiels. La différentiation étant une opération sensible au bruit, on lui associe généralement une opération de lissage pour réduire les effets du bruit.

Durant de nombreuses années, on a utilisé les masques de convolution comme opérateurs de différentiation, l'intérêt de ces filtres à réponse impulsionnelle finie résidant dans la rapidité des calculs. Désormais, la puissance des calculateurs permet l'utilisation de détecteurs plus élaborés et dits "optimaux".

2.6.2.1 Détecteur optimal

Un détecteur optimal est un opérateur devant répondre à des spécifications préalablement établies. Le choix du détecteur de contours le plus approprié pour l'application envisagée n'est pas trivial. Nous décrivons deux détecteurs basés sur des critères identiques mais d'étendues différentes. Puis nous détaillons l'implantation récursive du plus usité des deux, le filtre de Deriche [Der87]. Le lecteur intéressé par les détecteurs de contours pourra consulter :

- la thèse de Ziou [Zio91] qui définit un critère de sélection des détecteurs ainsi que le calcul automatique de leurs paramètres ;
- la thèse de Wan [Wan96] qui définit un détecteur de contours hyperbolique basé sur les critères de Canny [Can86] et dont les performances sont comparables à celles de l'opérateur de Deriche ;
- la thèse de Stoclin [Sto00] qui propose des critères discrets afin d'évaluer les performances des détecteurs de contours.

Les critères de Canny

Canny [Can86] définit les propriétés souhaitables pour un détecteur de contours performant. Pour cela, il propose les trois critères suivants :

1. La qualité de détection :

La détection sera de bonne qualité si l'influence du bruit sur l'image est faible. Ce critère correspond à maximiser le rapport signal sur bruit ;

2. La précision de localisation :

Ce critère permet de quantifier la précision de la position du contour détecté dans une image bruitée par rapport à sa position réelle ;

3. L'obtention d'une réponse unique pour un contour donné :

Ce critère sert à déterminer si le filtre est susceptible de fournir plusieurs réponses pour un seul contour présent dans l'image.

Dans [Can86], les points de contour correspondent aux maxima locaux du résultat de la convolution du signal avec un filtre. Soit $G(x)$ le signal correspondant à un contour de type échelon idéal et positionné en 0 :

$$G(x) = \begin{cases} A & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases} \quad (2.7)$$

La réponse $y(x)$ du filtre au signal en un point x_0 est donnée par la convolution entre le filtre $f(x)$ et le signal $G(x)$:

$$y(x_0) = \int_{-W}^{+W} G(x_0 - x)f(x)dx. \quad (2.8)$$

Notons que le détecteur $f(x)$ de Canny étant à support fini, il est uniquement défini dans l'intervalle $[-W, W]$.

Comme en pratique les images sont bruitées, Canny prend en compte l'existence d'un bruit gaussien de moyenne nulle et de variance η_0 qui se superpose au signal $G(x)$.

La qualité de détection

Afin de quantifier la qualité de la détection des contours, Canny considère le rapport signal sur bruit défini par le rapport entre la réponse du filtre au point de discontinuité du contour

($x_0 = 0$) et l'écart-type de la réponse du filtre au bruit :

$$\Sigma = \frac{|\int_{-W}^{+W} G(-x)f(x)dx|}{\eta_0 \sqrt{\int_{-W}^{+W} f^2(x)dx}}. \quad (2.9)$$

Comme les probabilités de ne pas détecter un vrai point de contour et de détecter de faux points de contour diminuent lorsque le rapport signal sur bruit augmente, le critère de qualité de détection correspond à la maximisation de ce rapport.

Pour l'échelon idéal défini par l'équation 2.7, on obtient :

$$\Sigma = \frac{A |\int_{-W}^0 f(x)dx|}{\eta_0 \sqrt{\int_{-W}^{+W} f^2(x)dx}}. \quad (2.10)$$

La précision de localisation

Ce critère indique la capacité du détecteur à extraire du signal bruité un contour qui est localisé aussi près que possible du contour réel. Pour ce critère, Canny propose de maximiser l'inverse de la variance de la distance entre le contour détecté et le contour réel :

$$\Lambda = \frac{|\int_{-W}^{+W} G'(-x)f'(x)dx|}{\eta_0 \sqrt{\int_{-W}^{+W} f'^2(x)dx}}. \quad (2.11)$$

Pour le contour de type échelon idéal, le critère de localisation devient :

$$\Lambda = \frac{A |f'(0)|}{\eta_0 \sqrt{\int_{-W}^{+W} f'^2(x)dx}}. \quad (2.12)$$

L'obtention d'une réponse unique

Le filtre $f(x)$ doit fournir un point unique en réponse à un seul contour. Afin de minimiser le nombre de réponses multiples, Canny propose d'utiliser la distance moyenne entre les passages par zéro de sa dérivée $f'(x)$ définie par :

$$x_{moy} = 2\Pi \sqrt{\frac{\int_{-W}^{+W} f'^2(x)dx}{\int_{-W}^{+W} f''^2(x)dx}}. \quad (2.13)$$

2.6.2.2 Le détecteur de Deriche

Canny a proposé une formulation d'un détecteur respectant ces critères. En modifiant l'étendue du filtre à l'infini, Deriche propose une formulation qui autorise une implantation récursive [Der87, MD89].

2.6.2.2.1 Le détecteur de Canny

Ayant choisi la dérivée première comme opérateur de différenciation, une contrainte d'antisymétrie du filtre $f(x)$ est imposée. Selon Canny, un "bon" détecteur de contours doit maximiser les trois critères préalablement définis. Afin de simplifier les calculs nécessaires à la détermination d'un détecteur optimal dans le cas de contours de type échelon, il a réduit la procédure de maximisation des deux premiers critères à la maximisation de leur produit $\Sigma\Lambda$. Le critère de réponses multiples est utilisé en tant que contrainte dans la procédure d'optimisation. Il impose que la distance maximale entre deux maxima adjacents de la réponse du filtre soit une constante, égale à k fois l'étendue W du filtre. La solution générale dans l'intervalle $[-W, 0]$ est la suivante :

$$f(x) = \begin{matrix} a_1 e^{\alpha_{can} x} \sin(\omega x) & + & a_2 e^{\alpha_{can} x} \cos(\omega x) & + \\ a_3 e^{-\alpha_{can} x} \sin(\omega x) & + & a_4 e^{-\alpha_{can} x} \cos(\omega x) & + & c \end{matrix} \quad (2.14)$$

où α_{can} et ω sont des constantes réelles. Les coefficients a_1, a_2, a_3, a_4 et c peuvent être déterminés grâce aux conditions limites suivantes :

$$f(0) = 0, \quad f(W) = 0, \quad f'(0) = S, \quad f'(W) = 0, \quad \text{avec } S \text{ constant.} \quad (2.15)$$

Comme $f(x)$ est antisymétrique, on peut étendre cette solution à l'intervalle $[-W, W]$. Canny propose finalement comme détecteur la dérivée première d'une gaussienne qui est une bonne approximation du détecteur initial (cf. équation 2.14) :

$$f(x) = E x e^{-\alpha_{can}^2 x^2} \quad (2.16)$$

où E est une constante permettant de normaliser le filtre et α_{can} un paramètre permettant de régler l'étendue efficace du filtre. Cette approximation a permis d'étendre le filtre de Canny au domaine bidimensionnel.

2.6.2.2.2 Le détecteur optimal de contours de Deriche

En étendant le filtre de Canny à l'infini, c'est à dire $W \rightarrow \infty$, Deriche obtient une sinusoïde amortie comme réponse impulsionnelle du filtre de détection de contour :

$$f(x) = C e^{-\alpha_{der}|x|} \sin(\omega x) \quad (2.17)$$

où ω est une constante positive. C est un coefficient de normalisation permettant d'obtenir une réponse unitaire pour un échelon unitaire. Le paramètre α_{der} définit l'étendue de la largeur

efficace du filtre, ce qui correspond également au compromis entre la détection et la localisation. Lors de ses travaux, Deriche a constaté qu'il obtenait de meilleurs résultats lorsque $\omega \simeq 0$, ce qui implique $\sin(\omega x) \simeq \omega x$. Le détecteur optimal devient alors :

$$d(x) = C\omega x e^{-\alpha_{der}|x|} \quad (2.18)$$

L'opérateur de lissage

Afin de limiter l'influence du bruit amplifié par l'opérateur de différentiation, on peut faire appel à différents filtres. Deriche propose d'utiliser un filtre de type lissage correspondant à l'intégrale de l'opérateur de détection (cf. équation 2.18). La définition de ce filtre est alors :

$$l(x) = B(\alpha_{der} |x| + 1)e^{-\alpha_{der}|x|} \quad (2.19)$$

où B , une constante réelle, représente un coefficient de normalisation et où α_{der} est un paramètre permettant de régler l'étendue efficace de l'opérateur de lissage.

Implantation récursive

Pour une implantation sur ordinateur, il convient de redéfinir l'équation du filtre dans le domaine numérique en utilisant la transformée en Z .

L'implantation récursive du détecteur optimal $d(x)$ (cf. équation 2.18) est réalisée par deux filtres récursifs opérant dans des directions opposées. Soient $x(n)$ le signal d'entrée et $y(n)$ le signal de sortie, tel que $y(n) = y^+(n) - y^-(n)$ pour $n = 1, \dots, N$ avec N correspondant au nombre d'échantillons du signal. Les expressions de y^+ et y^- sont données ci-après :

$$\begin{aligned} y^+(n) &= ax(n-1) - b_1 y^+(n-1) - b_2 y^+(n-2) \text{ pour } n = 1, \dots, N \\ y^-(n) &= ax(n+1) - b_1 y^-(n+1) - b_2 y^-(n+2) \text{ pour } n = N, \dots, 1 \end{aligned}$$

Lors de l'initialisation, les valeurs de $y^+(-1)$, $y^+(0)$, $y^-(N+1)$ et $y^-(N+2)$ sont mises à 0. Les valeurs des constantes a , b_1 et b_2 sont les suivantes :

$$a = De^{-\alpha_{der}}, \quad b_1 = -2e^{-\alpha_{der}}, \quad b_2 = e^{-2\alpha_{der}}$$

L'implantation récursive du filtre de lissage $l(x)$ (cf. équation 2.19) est réalisée de la même manière, avec les mêmes notations que celles utilisées ci-dessus :

$$\begin{aligned} y^+(n) &= a_0 x(n) + a_1 x(n-1) - b_1 y^+(n-1) - b_2 y^+(n-2) \text{ pour } n = 1, \dots, N \\ y^-(n) &= a_2 x(n+1) + a_3 x(n+2) - b_1 y^-(n+1) - b_2 y^-(n+2) \text{ pour } n = N, \dots, 1 \end{aligned}$$

Les valeurs des constantes a_0, a_1, a_2, a_3, b_1 et b_2 sont les suivantes :

$$a_0 = s, \quad a_1 = s(\alpha_{der} - 1)e^{-\alpha_{der}}, \quad a_2 = a_1 - sb_1,$$

$$a_3 = -sb_2, \quad b_1 = -2e^{-\alpha_{der}}, \quad b_2 = e^{-2\alpha_{der}},$$

$$s = \frac{(1 - e^{-\alpha_{der}})^2}{1 + 2\alpha_{der}e^{-\alpha_{der}} - e^{-2\alpha_{der}}}.$$

2.6.2.3 Extraction des points de contour

La dernière étape de l'extraction des primitives consiste à ne retenir que les points de contour pertinents, c'est à dire ceux censés correspondre aux contours d'un objet. Pour cela, on ne sélectionne dans l'image module de la dérivée que les extrema locaux supérieurs à un seuil. Nous décrirons plus loin dans ce paragraphe une méthode permettant de déterminer automatiquement ce seuil. Notons que, dans le cadre de la recherche des contours, une analyse de connexité des points de contour est souvent requise. C'est notamment le cas pour le filtrage proposé par Canny [Can86] qui effectue un seuillage par hystérésis. Quoi que souhaitable, cette méthode ne peut être applicable dans notre cas car nous travaillons dans un contexte monodimensionnel, en opérant ligne par ligne.

Lors de ses expérimentations, Burie a constaté que la convolution d'une image avec un filtre de détection de contour pouvait donner naissance à plusieurs extrema locaux pour un contour unique. Afin de sélectionner les extrema significatifs, il a développé une nouvelle technique d'extraction de primitives, baptisée *extraction par intervalle*. Celle-ci procède tout d'abord à l'élimination des extréma générés par le bruit par un simple seuillage de la réponse de l'opérateur de Deriche. Puis il divise le signal dérivé en une suite d'intervalles adjacents dans lesquels ce signal conserve un signe constant. L'analyse est faite du début à la fin de l'image et un nouvel intervalle est créé chaque fois que le signal dérivé passe par zéro. Un seul extremum, quand il existe, est ensuite sélectionné dans chaque intervalle : il s'agit de l'extremum pour lequel l'amplitude de la dérivée est la plus forte (cf. figure 2.8). Trois attributs sont associés à chacune de ces primitives de type points de contour. Il s'agit de :

- leur position dans l'image-ligne ;
- le signe de la dérivée en ce point ;
- l'amplitude de la dérivée en ce point.

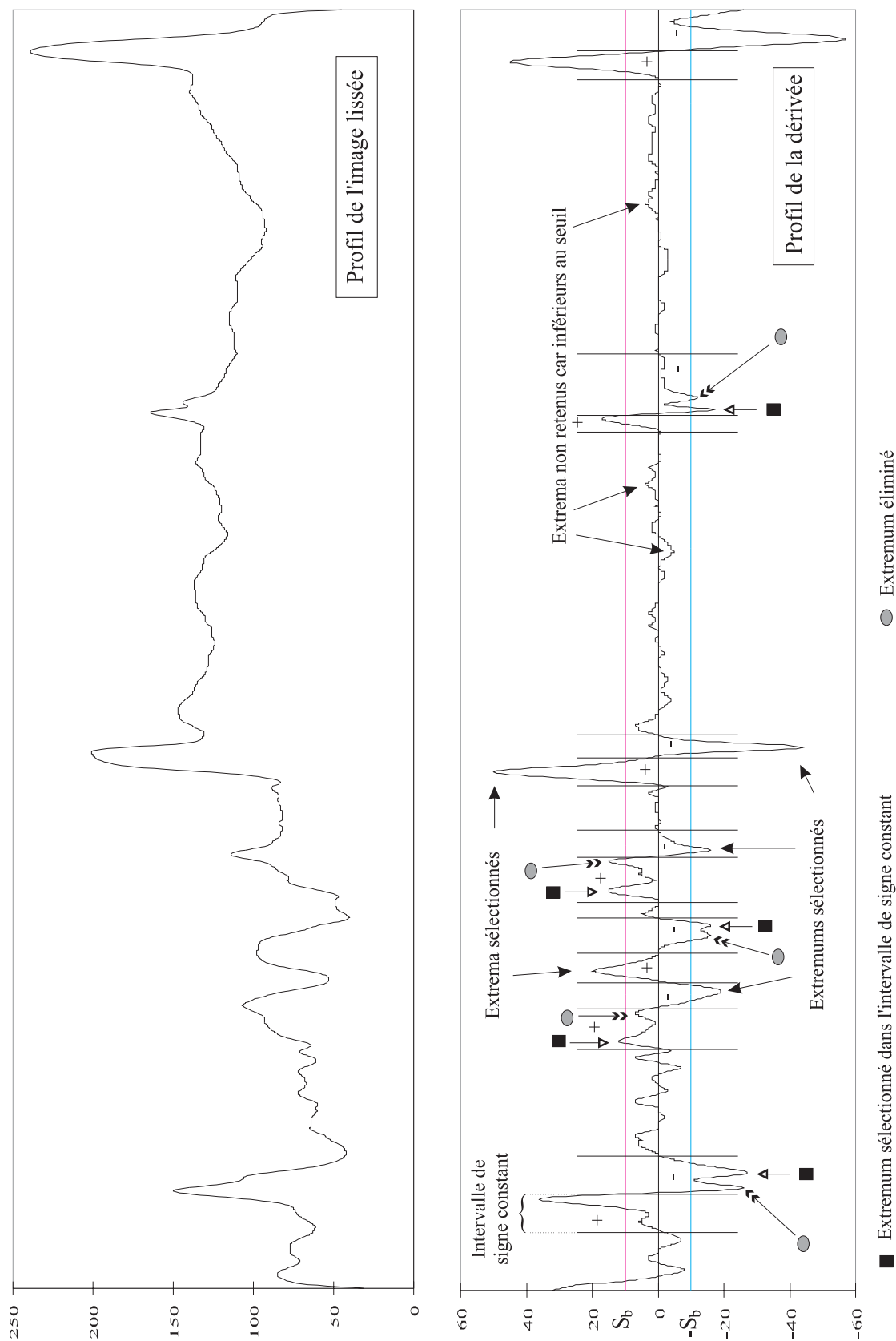


FIG. 2.8 : Extraction des extremums par intervalles de signe constant.

2.6.3 Conclusion

Dans ce paragraphe nous avons défini le type de primitives que nous allons utiliser. Ce choix a été imposé par la stratégie que nous avons adoptée, qui consiste à travailler sur les droites épipolaires. Ensuite, nous avons décrit le détecteur de points de contour de Deriche, optimal au sens des critères de Canny. Son implantation récursive permet un gain de temps important et ce, quelle que soit l'étendue efficace de ce filtre. Le réglage de ce filtre s'effectue par le biais d'un unique paramètre α_{der} qui est généralement ajusté par expérience. Ensuite, afin de ne conserver que les extrema significatifs, nous utilisons la technique *d'extraction par intervalle* développée par Burie.

La procédure de détection des points de contour appliquée sur des images stéréoscopiques réelles ou de synthèse fournit deux listes de points de contour, une pour l'image gauche et l'autre pour l'image droite. Chacune de ces listes est composée des points de contour non horizontaux extraits des images stéréoscopiques respectives. Trois attributs sont associés à chacun de ces points, à savoir :

- leur position dans l'image ;
- le signe de la dérivée en ce point ;
- l'amplitude de la dérivée en ce point.

Une fois que la phase d'extraction des primitives est terminée, la procédure de mise en correspondance des primitives extraites des deux images peut être activée. Mais une primitive d'une image peut être appariée à plusieurs primitives de l'autre image. Afin d'établir des appariements corrects, l'utilisation de critères de ressemblance et de contraintes stéréoscopiques paraît indispensable. Dans le paragraphe suivant, nous présentons ces contraintes et ces critères. Ils sont utilisés pour réduire l'espace de recherche afin d'établir des appariements sans ambiguïté.

2.7 Critères de ressemblance et contraintes

La mise en correspondance des primitives nécessite de définir des critères qui permettent de fournir des informations sur leur ressemblance. Cependant, l'utilisation des critères de ressemblance est souvent insuffisante pour décider si un appariement est correct ou non. En effet, une primitive d'une image peut ressembler à plusieurs primitives de l'autre image. Pour résoudre ce problème d'ambiguïté, on fait appel à un certain nombre de contraintes.

2.7.1 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 de deux primitives est que les attributs qui leurs sont attachés doivent être semblables si ces deux primitives se correspondent. L'idée de base consiste à apparier une primitive en la comparant à ses correspondantes 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 qui permet 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. Étant donnée une primitive, la recherche de son homologue consiste à sélectionner parmi les appariements possibles celui qui optimise le critère.

Selon le type de primitives utilisé, les critères de ressemblance peuvent être de deux natures : photométriques ou géométriques. Le premier type de critères est utilisé essentiellement lorsqu'il s'agit d'apparier des primitives de type point [HM93, Fa93]. Le second type s'applique aux primitives de plus haut niveau (contours, régions, etc.) caractérisées par des attributs géométriques (longueurs des segments, angles entre les segments, surfaces et périmètres des régions, etc.). Ce type est inutilisable dans le cadre de cette thèse car les primitives que nous cherchons à mettre en correspondance sont les points de contour détectés dans les images stéréoscopiques.

2.7.2 Contraintes stéréoscopiques

L'utilisation des critères de ressemblance n'est en général pas suffisante pour établir des appariements sans ambiguïté, une primitive de l'image gauche pouvant ressembler à plusieurs primitives de l'image droite et réciproquement. Comment peut-on donc décider que l'appariement d'une paire de primitives, gauche et droite, est correct ou non ? Pour répondre à cette question, on considère un certain nombre de contraintes. Celles-ci sont de deux types : les contraintes locales et les contraintes globales. Les contraintes locales sont applicables à chaque couple de primitives candidates à l'appariement, sans faire intervenir les autres couples. La contrainte de position et la contrainte de signe de la dérivée font partie de ce type des contraintes locales. Contrairement aux contraintes locales, les contraintes globales sont applicables à tous les couples de primitives candidats à l'appariement et permettent de valider la compatibilité entre les appariements satisfaisant le premier type de contraintes. Parmi les principales contraintes

de ce type, nous pouvons citer les contraintes d'ordre, d'unicité et de continuité de la disparité. L'ensemble de ces contraintes est détaillé dans le chapitre suivant.

2.8 Mise en correspondance

Dans un processus stéréoscopique, la tâche la plus difficile et la plus délicate est celle concernant la mise en correspondance des primitives extraites de deux images stéréo. Il n'existe pas, jusqu'à présent, une méthode assez générale pour résoudre ce problème et qui puisse s'appliquer à n'importe quelle paire d'images stéréoscopiques sans tenir compte de leur type, 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 au sein d'un ensemble fini. 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 d'erreur dû à de faux appariements, qui doit être aussi faible que possible.

2.8.1 Méthodes corrélatives

Les méthodes corrélatives sont utilisées généralement pour mettre en correspondance les primitives de type point [KO94]. La recherche des appariements est basée sur un critère de ressemblance photométrique : elle consiste à considérer une fenêtre autour du point que l'on veut appairer, puis à chercher dans l'autre image le point dont la fenêtre, de même taille que la précédente, est la mieux corrélée avec la première fenêtre. Avant chaque mesure de ressemblance, la contrainte épipolaire est utilisée afin de réduire le nombre de candidats à l'appariement. Après optimisation du critère de ressemblance sur l'ensemble des appariements possibles, des contraintes globales (ordre, unicité, continuité de la disparité) sont ensuite utilisées pour lever certaines ambiguïtés.

2.8.2 Méthodes basées sur la programmation dynamique

Le problème de l'appariement peut être posé sous forme de recherche d'un chemin optimal dans un graphe [LL94a, OBMD94]. 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 une matrice à deux dimensions, où chaque dimension est définie par les primitives de chacune des droites

épipolaires gauche et droite. A chaque sommet du graphe correspond un élément de la matrice 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 est constitué d'une séquence de paires de primitives gauche et droite. Le coût d'un chemin est calculé en fonction de celui affecté à chaque paire qui le constitue. L'objectif est donc de trouver un chemin optimal, c'est à dire de coût minimum, dans ce graphe. La nature de la recherche d'un chemin optimal, qui est liée au 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 n'intervient qu'une seule fois (contrainte d'unicité).

2.8.3 Appariement par prédiction et vérification d'hypothèses

Le processus de mise en correspondance par prédiction et vérification d'hypothèses se décompose en trois étapes indépendantes [Aya89, KHK90, Ben94] :

1. *Prédiction d'hypothèses.* Après avoir extrait de chaque image les primitives à appairier, la première étape consiste à établir un ensemble d'appariements à partir de contraintes locales exclusivement, telles que la contrainte épipolaire et la contrainte d'orientation, et de critères de ressemblance géométriques et/ou photométriques.
2. *Propagation d'hypothèses.* Après avoir généré un certain nombre d'appariements hypothétiques lors de la phase de prédiction, la phase de propagation a pour but d'étendre le nombre d'appariements. Il s'agit de propager chacune des hypothèses aux voisines des primitives appariées. Supposons qu'on émette l'hypothèse d'appariement des primitives (P_g, P_d) . La propagation de cette hypothèse consiste à chercher pour les P_{gi} au voisinage de P_g les appariements à des primitives P_{dj} au voisinage de P_d , en utilisant les contraintes de continuité de la disparité et d'unicité par exemple. Les nouvelles hypothèses ainsi générées sont, elles aussi, propagées de manière récursive.
3. *Vérification d'hypothèses.* Cette dernière phase est destinée à comparer les appariements obtenus par propagation des hypothèses prédites. Les appariements retenus sont ceux qui correspondent au plus grand nombre d'hypothèses compatibles vis-à-vis de la contrainte de continuité de la disparité par exemple.

2.8.4 Méthodes séquentielles

Les algorithmes séquentiels de mise en correspondance sont définis selon les principes suivants :

- l'appariement des primitives s'effectue *séquentiellement* en explorant les droites épipolaires conjuguées du début à la fin ;
- lorsqu'un couple de primitives est apparié, il sert de base pour trouver l'appariement suivant.

Plus précisément, il s'agit d'analyser les listes de primitives $G(i)$ et $D(j)$ situées sur les droites épipolaires conjuguées en faisant croître les indices i et j . En utilisant la représentation graphique du plan de mise en correspondance proposée par Otha [OK85], l'appariement des primitives peut être considéré comme la recherche d'un chemin dont chaque étape est constituée d'un nœud correspondant à une paire stéréoscopique. Initialement, la procédure débute dans le coin supérieur gauche du plan, l'objectif étant d'atteindre le coin inférieur droit.

Ceci justifie la nécessité de définir le premier et le dernier pixel des droites épipolaires comme étant des points de contour. Le second principe consiste à utiliser la dernière paire valide pour trouver le prochain couple à appairer.

Le principe de mise en correspondance peut alors s'énoncer de la manière suivante : à partir de la paire valide courante $\{G(i), D(j)\}$, on peut déterminer à l'aide d'une mesure de similarité, le prochain appariement correct défini par la paire $\{G(i+k), D(j+k')\}$ avec $1 \leq k \leq N_G - i$ et $1 \leq k' \leq N_D - j$. Le lecteur souhaitant obtenir plus d'informations sur cette méthode et sur la mesure de similarité définie par Burie pourra se reporter à son manuscrit [Bur95].

Ayant constaté quelques faux appariements dans ses séquences d'images, Burie a proposé de rendre plus robuste sa méthode de mise en correspondance en procédant à une exploration bi-directionnelle des listes de primitives gauche et droite.

2.8.4.1 Méthode bidirectionnelle

Cette exploration bidirectionnelle est composée de deux passes : la première, du début vers la fin des images, s'appelle *l'analyse progressive*. La seconde passe est effectuée de la fin vers le début, il s'agit de *l'analyse rétrograde*. Chaque analyse fournit une liste de paires valides. Une comparaison des deux listes permet ensuite d'éliminer certaines erreurs d'appariement.

Les paires valides qui sont retenues doivent respecter les contraintes d'ordre et d'unicité. On peut rencontrer trois types de situations :

1. *une paire $\{G(i), D(j)\}$ apparaît dans les deux listes.* Dans ce cas, la paire a été appariée deux fois. Elle est donc rangée dans la liste définitive ;
2. *un point de contour d'une image est apparié à deux points de contour différents de l'autre image.* La contrainte d'unicité impose de faire un choix parmi ces deux appariements. On conserve alors le couple qui respecte la contrainte d'ordre avec le couple précédemment stocké dans la liste définitive. Si les deux couples vérifient la contrainte d'ordre, on garde celui qui présente la mesure de similarité la plus forte ;
3. *une paire $\{G(i), D(j)\}$ n'est présente que dans une seule liste.* Dans ce cas, la paire est conservée si elle vérifie la contrainte d'ordre avec le dernier couple stocké dans la liste définitive.

Cette méthode est lourde à mettre en œuvre car elle nécessite d'effectuer deux passes ainsi qu'une comparaison des résultats fournis par les deux analyses. De plus, le point faible de l'algorithme réside dans le fait que toute erreur d'appariement se propage dans le plan de mise en correspondance. Burie propose alors d'apparier hiérarchiquement les points de contour en utilisant la notion de niveau de recherche.

2.8.4.2 Méthode hiérarchique

L'idée principale de cette méthode est de déterminer les couples de points de contour de référence à partir desquels sera effectuée la mise en correspondance des autres points de contour. Pour un niveau de recherche donné, on commence par sélectionner dans chaque image les points de contour les plus significatifs qui correspondent aux valeurs les plus élevées de l'amplitude de la dérivée. On applique alors la méthode de mise en correspondance à ces points. Les appariements réalisés servent de base pour rechercher les appariements des primitives moins significatives qui correspondent à des amplitudes de la dérivée moins élevées.

2.8.5 Méthodes basées sur les réseaux de neurones

Dès 1976, Marr et Poggio [MP76] ont développé un algorithme itératif faisant coopérer les contraintes de continuité de la disparité et d'unicité. Ils ont montré que cet algorithme pouvait

être implanté en utilisant un réseau de neurones sélectifs à différentes combinaisons de coordonnées spatiales et de disparités et dont les poids de connexions sont calculés au préalable. Les neurones sont connectés de telle sorte qu'il y ait 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. [KBL93] ont considéré, dans un premier temps, le problème de mise en correspondance comme une application d'un ensemble d'appariements initiaux, vérifiant la contrainte épipolaire, vers l'ensemble des appariements corrects. Les primitives utilisées sont les pixels des deux images stéréoscopiques. Cette application est ensuite implantée sur un réseau de neurones multicouche (couche d'entrée, couche cachée, couche de sortie) en effectuant l'apprentissage par rétropropagation du gradient. Avec cette technique, seule la contrainte épipolaire est imposée explicitement. Les autres contraintes (unicité et continuité de la disparité) sont codées par le réseau au cours de l'apprentissage.

Plus récemment, Cruz et al. ont proposé d'apparier des segments de droite en utilisant un réseau de neurones multicouche basé sur une technique d'auto-organisation [CPA95]. Tout d'abord, le problème est ramené à un problème de classification. Il s'agit de classer toutes les paires de primitives, gauches et droites, dans deux classes : "appariement correct" ou "appariement incorrect". Un apprentissage compétitif non supervisé est appliqué au réseau pour assigner chaque appariement possible à l'une de ces deux classes.

Une autre approche consiste à considérer le problème d'appariement comme un problème d'optimisation. Dans un premier temps, une fonction de coût, représentant les contraintes du problème (unicité, continuité de la disparité, ordre, etc.), est construite de telle façon que ses minima correspondent à des appariements acceptables. Un réseau de neurones de Hopfield, entièrement connecté, est ensuite utilisé pour minimiser cette fonction de coût [NC92, LSH94, Rui97].

2.8.6 Approche par optimisation utilisant le recuit simulé

L'inconvénient majeur de l'approche par réseaux de Hopfield est que le processus de minimisation associé à ce type de réseau est local, c'est à dire qu'il tend à proposer une solution qui est le premier minimum local rencontré. Pour pallier cette limitation, certains auteurs ont proposé d'utiliser l'algorithme du recuit simulé [Kir84]. L'idée de base consiste à considérer

le système à optimiser comme un système physique défini par ses états et par son énergie. Le minimum absolu de cette énergie est atteint en simulant un processus de recuit en partant d'une température élevée pour terminer à une température voisine de zéro.

La qualité des résultats produits par l'algorithme du recuit simulé est satisfaisante car le processus du recuit force l'état du système dans des régions de basse énergie tout en lui évitant d'être piégé dans des états correspondant à des minima locaux où l'énergie est élevée. De plus, son implantation sur des architectures neuronales de type réseau de Hopfield est naturelle : il suffit d'associer à chaque état élémentaire du système la sortie d'un neurone et d'identifier l'énergie à minimiser à celle du réseau [Rui97].

Le lecteur intéressé par les méthodes d'appariement stéréoscopique par recuit simulé pourra consulter [VG91, MTM93, GL93].

2.8.7 Approche par algorithmes génétiques

Un algorithme génétique est une procédure itérative qui utilise une population de solutions potentielles à un problème donné. Chaque solution est codée sous forme d'un chromosome composé d'une chaîne de gènes qui peuvent prendre une de différentes valeurs appelées *allèles*. Chaque itération consiste en trois étapes : évaluation, sélection et reproduction (croisement et mutation).

Récemment, plusieurs auteurs se sont penchés sur l'utilisation des AGs pour résoudre le problème de mise en correspondance en stéréovision. Les résultats obtenus sont encourageants et confirment l'efficacité des AGs quand ils sont appliqués dans ce domaine. En effet, les AGs peuvent traiter des problèmes pour lesquels les solutions potentielles sont situées dans un espace de grande dimension, contrairement aux méthodes standard reposant sur une exploration systématique. De plus, avec un algorithme génétique, la recherche d'une solution optimale à un problème ne nécessite aucune connaissance a priori sur la répartition des solutions dans l'espace.

D'une manière générale, les AGs sont différents des méthodes classiques, indépendamment du problème traité. Leurs spécificités viennent de :

- La manipulation directe d'un codage des paramètres, et non les paramètres eux-mêmes.
- L'exploration au moyen d'une population de points, au lieu d'un point unique.
- L'exploration uniquement à partir des données des problèmes. En pratique, seules les

valeurs de la fonction d'évaluation (fonction à optimiser) sont utilisées, sans faire appel à sa dérivée ou d'autres connaissances auxiliaires.

- L'utilisation de règles de transition stochastiques permettant une exploration globale et efficace de l'espace de recherche des solutions.

Pour illustrer l'intérêt porté par les chercheurs aux algorithmes génétiques dans le domaine de la stéréovision et bien situer notre contribution dans ce cadre, nous allons décrire dans les deux sous-paragraphes suivants quelques travaux récemment publiés.

2.8.8 Algorithme des mouches dynamiques (approche parisienne)

L'algorithme des mouches est un algorithme rapide d'évolution artificielle conçu pour l'exploration de l'espace des paramètres dans certaines applications de reconnaissance des formes. Louchet et al. [LGLB01] l'ont utilisé pour le guidage d'un robot en temps réel. Dans cette application, la population est un ensemble de particules ("mouches") situées dans l'espace, et qui constitue une représentation collective ("approche parisienne"⁶[CLRS99, Lou00]) de la scène en trois dimensions. En d'autres termes, l'algorithme exploite un couple d'images stéréoscopiques en utilisant des points $3D$ ("mouches") comme individus et fait évoluer cette population grâce à une fonction d'évaluation telle que les mouches convergent vers les surfaces visibles des objets de la scène. En effet, le chromosome n'est que le triplet (x, y, z) qui contient les coordonnées de la mouche et la fonction d'évaluation caractérise la ressemblance des voisinages des projections de la mouche dans les deux images. Si la mouche se trouve à la surface d'un objet opaque, alors les pixels correspondant dans les deux images doivent normalement avoir le même niveau de gris⁷ et des voisinages semblables. Inversement, si la mouche n'est pas sur la surface d'un objet et que ce dernier est suffisamment hétérogène, la ressemblance entre les voisinages de ses projections sera faible (*cf.* figure 2.9).

Pour pouvoir traiter une séquence d'images stéréoscopiques, Louchet et al. ont étendu l'algorithme décrit ce-dessus en faisant évoluer en continu la population au cours de la séquence d'images. Cela signifie qu'au cours de l'évolution, la fonction d'évaluation n'est plus statique mais dépend des images courantes. En effet, la population est initialisée par le nuage de mouches

⁶L'approche parisienne de l'évolution artificielle considère chaque individu comme une partie de la solution du problème : la solution est représentée par l'ensemble de la population, ou une large fraction de celle-ci.

⁷Ou les mêmes composantes chromatiques dans le cas d'images en couleurs. Cette propriété est bien vérifiée avec les surfaces mates (lambertiennes) où la diffusion de la lumière est isotrope. Les réflexions spéculaires peuvent créer des objets virtuels (au sens de l'optique géométrique) et fausser l'interprétation $3D$, quel que soit l'algorithme de traitement stéréo.

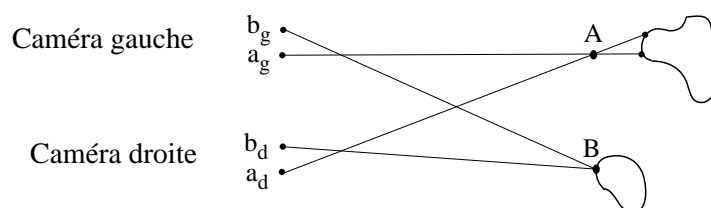


FIG. 2.9 : Les pixels b_g et b_d , projections de la mouche B , sont fortement corrélés, à la différence des pixels a_g et a_d , projections de A , qui reçoivent leur illumination de deux points physiques différents de la scène.

obtenu sur le couple stéréoscopique précédent, ce qui peut-être vu comme une forme de mémorisation de la représentation de l'espace. Pour que leur algorithme étendu puisse s'adapter au cas des mouvements rapides, Louchet et al. introduisent de chromosomes étendus qui contiennent les composantes de positions et de vitesse des mouches dans le repère mobile lié à la caméra de référence. En d'autres termes, le chromosome associé à une mouche est le 6-uplet $(x, y, z, x^\bullet, y^\bullet, z^\bullet)$.

Les résultats obtenus grâce à l'approche parisienne dans le domaine de la robotique mobile sont très encourageants car elle permet de partager la représentation de la scène en de très nombreuses primitives très simples. La fonction d'évaluation est donc elle-même simple et l'algorithme est rapide.

2.8.9 Évaluation de la carte dense de disparité par algorithmes génétiques

La plupart des méthodes d'appariement par corrélation mesurent la ressemblance entre deux pixels en calculant un score de corrélation (ou critère de similarité) déterminé sur leur voisinage. Le pixel et son voisinage sont appelés *fenêtre* ou *domaine de corrélation*. Le choix de la taille des fenêtres est crucial. Trop petites, elles ne contiennent pas suffisamment d'information pour permettre un appariement fiable. Trop grandes, elles peuvent contenir des éléments appartenant à des objets différents situés à des distances différentes. Afin de pallier ce défaut, Kanade [KO94] propose une solution en utilisant une fenêtre adaptative de taille variable.

Pour déterminer la taille optimale des fenêtres permettant la meilleure mise en correspondance, Saito et Mori [SM95] évaluent les cartes de disparités générées pour différentes tailles de fenêtres. Pour un pixel donné, ils obtiennent donc plusieurs valeurs de disparité. Ensuite, par le biais des algorithmes génétiques, ils sélectionnent celle qui minimise une fonction de coût. L'algorithme se décompose en deux étapes :

- plusieurs cartes de disparités sont obtenues en utilisant le calcul de la corrélation pour

différentes tailles de fenêtres. On obtient ainsi plusieurs valeurs de disparité pour chaque pixel ;

- ensuite, l'application des algorithmes génétiques permet de sélectionner la disparité optimale pour chaque pixel, en recherchant le minimum d'une fonction de coût.

Saito et Mori relèvent que les erreurs de compatibilité et de continuité de la carte de disparités sont moindre que celles fournies par un algorithme utilisant des tailles de fenêtres fixes.

Gong et Yang [GY02] considèrent que l'approche développée par Saito et Mori est limitée du fait que la vraie valeur de disparité est supposée être l'une des valeurs candidates calculées grâce aux différentes fenêtres utilisées. Or ceci n'est peut être pas toujours valable quand il s'agit d'images bruitées ou quand une partie de la scène observée est de couleur uniforme et manque d'objets à apparier. Pour améliorer la précision de la carte dense de disparité, Gong et Yang suppriment, tout d'abord, de cette dernière les mauvais appariements causés par les occlusions en utilisant plusieurs images stéréoscopiques. Ensuite, ils utilisent un algorithme génétique afin d'éliminer les mauvais appariements en minimisant les erreurs de compatibilité des points appariés et de continuité de la carte de disparités. Cette approche fait partie des méthodes appelées *global intensity-based methods* qui tentent d'introduire des contraintes globales dans le processus d'appariement ⁸. L'idée de base de ces méthodes consiste à construire un espace tridimensionnel (U, V, D) , appelé *espace de disparité*, dans lequel la valeur d'un voxel $P_{(u,v,d)}$ indique la probabilité que le pixel (u, v) ait la disparité d . La méthode de mise en correspondance utilisée détermine alors la surface dans cet espace qui satisfait les deux contraintes suivantes : 1- la surface est lisse ; 2- elle passe par les voxels ayant des valeurs de probabilités élevées. Cette surface n'est rien d'autre que la carte de disparité recherchée.

L'approche proposée par Gong et Yang peut-être résumée comme suit : tout d'abord, un espace de disparité tridimensionnel est peuplé avec les valeurs de similarité des pixels ayant les mêmes coordonnées dans les images stéréoscopiques binoculaires⁹. Quand il s'agit de plusieurs images stéréoscopiques, l'espace est rempli par des valeurs de la fonction de détection d'occlusion proposée par Satoh et Ohta [SO94, SO95] ¹⁰.

En ce qui concerne le codage génétique de cette approche, un chromosome qui code une

⁸C'est-à-dire que la disparité d'un pixel est influencée par les disparités des pixels voisins.

⁹N'importe quelle fonction produisant de valeurs élevées quand les voisinages des pixels en question ont de couleurs proches les uns des autres, peut être utilisée.

¹⁰Cette fonction est utilisée pour détecter les mauvais appariements dus aux occlusions. Elle a été proposée dans leur approche connue sous le nom de *Stereo by Eye Array algorithm*.

carte de disparité possible n'est qu'un arbre quaternaire qui fournit un schéma multi-résolution des images stéréoscopiques. Étant donné que les nœuds à différents niveaux d'un arbre quaternaire contiennent un nombre différent de pixels, la sélection des nœuds de différents niveaux de résolution produit un effet semblable à celui de l'ajustement de la taille de la fenêtre de corrélation aux différents endroits de l'image. Les arbres quaternaires utilisés pour représenter les cartes de disparité doivent satisfaire les contraintes suivantes :

- Chaque feuille (un nœud sans enfant) k d'un arbre quaternaire est associée à une valeur de disparité x , qui impose que les valeurs de disparité de tous les pixels couverts par k soit x . Une feuille k couvre un pixel p si elle contient le pixel p .
- Un nœud intérieur dans un arbre quaternaire ne peut pas avoir tous ses descendants assignés à la même valeur de disparité. Sinon, tous les descendants de ce nœud doivent être supprimés et le nœud lui-même est considéré comme une feuille.

Afin d'évaluer les cartes de disparité codées sous forme de chromosomes, une fonction d'énergie est définie en se basant sur les champs aléatoires de Markov ¹¹. Un AG est ensuite utilisé afin d'extraire la meilleure surface de l'espace de disparité par rapport à la fonction d'évaluation.

Globalement, les cartes de disparité fournies par cette approche sont meilleures que celles générées par d'autres méthodes. Cependant, les mesures numériques effectuées pour comparer les résultats de cette approche avec d'autres méthodes montrent que, lorsqu'il s'agit de la carte de disparité ayant le taux le plus élevé d'appariements corrects, celle-ci ne correspond pas à celle qui est meilleure visuellement. Le développement d'autres mesures plus appropriées mériterait davantage de recherche.

2.9 Conclusion

Un des objectifs de la vision artificielle est de percevoir le relief d'un environnement à partir d'au moins deux images. Les recherches dans ce domaine ont permis de développer de nombreuses approches. La stéréovision constitue l'une des méthodes les plus répandues et les plus utilisées pour aborder ce problème. Il s'agit de retrouver la géométrie spatiale d'une scène à partir de plusieurs images de cette même scène, prises sous des angles différents.

Les différentes méthodes développées pour l'appariement d'images en stéréovision peuvent-

¹¹Les champs de Markov aléatoires sont utilisés pour modéliser l'interaction entre les pixels voisins.

être regroupées en deux catégories : *méthodes denses* et *méthodes non-denses*. Dans cette thèse, nous nous intéressons à la deuxième catégorie de méthodes. Pour valider nos algorithmes, nous nous servons des images issues des stéréoscopes développées au laboratoire d'Automatique I^3D pour la détection d'obstacles. La première étape qui consiste à extraire les primitives à appairer est appliquée sur les images stéréoscopiques selon les 3 étapes suivantes :

- Une étape de lissage destinée à réduire l'effet du bruit présent dans les images-ligne qui ne sont rien d'autre que les droites épipolaires conjuguées.
- Une étape de différentiation, appliquée à chacune des images-ligne lissées. Cette étape est destinée à mettre en évidence les points de contour, c'est-à-dire les points correspondant aux extrema locaux des images-ligne dérivées.
- Une étape d'extraction destinée à ne conserver que les extrema locaux significatifs correspondant aux points de contour pertinents des deux images-ligne.

À ce stade du traitement, nous disposons de deux listes de points de contour auxquels sont associés les éléments suivants :

- La position du point de contour dans l'image-ligne correspondante ;
- Le signe de la dérivée en ce point ;
- La valeur de l'amplitude de la dérivée en ce point.

Dans le chapitre suivant, nous présentons l'approche génétique que nous avons développée pour l'appariement des points de contour extraits à partir des images-ligne stéréoscopiques.

Chapitre 3

Approche globale pour l'appariement d'images stéréoscopiques : Mise en œuvre par algorithmes génétiques

3.1 Introduction

La mise en correspondance des primitives extraites d'images stéréoscopiques est la phase la plus difficile et la plus délicate en stéréovision. Pour la résoudre, nous proposons une approche globale basée sur l'utilisation des algorithmes génétiques. L'idée principale de cette approche est de ramener le problème d'appariement d'images stéréoscopiques à un problème d'optimisation. Pour ce faire, il faut tout d'abord trouver un codage approprié permettant de représenter les solutions du problème sous forme de chromosomes. Ensuite, il est nécessaire de définir une fonction d'évaluation appelée également "fonction objectif" permettant de quantifier l'écart entre la solution recherchée et les chromosomes. Cette fonction permet d'avantager ou de désavantager une solution vis-à-vis des règles de sélection et de reproduction afin d'améliorer l'adaptation d'un ensemble de chromosomes par rapport à l'objectif fixé. Cet ensemble de chromosomes est appelé une population. Une fois le codage génétique et la fonction d'évaluation définis, l'algorithme d'optimisation explore l'espace de recherche des solutions grâce aux opérations de sélection, de croisement et de mutation des chromosomes. Ces opérations sont répétées jusqu'à ce qu'un critère d'arrêt soit atteint.

Dans ce chapitre, nous allons rappeler brièvement les principes de base de la méthode d'optimisation par algorithme génétique. Nous présentons notamment les techniques de sélection, de croisement et de mutation. Puis nous décrivons le codage génétique proposé pour résoudre le problème de mise en correspondance d'images stéréoscopiques. Ensuite, nous présentons la

fonction d'évaluation construite à partir des contraintes stéréoscopiques de telle sorte que ses extrema (minima ou maxima) correspondent à des appariements corrects. Avant d'appliquer cette approche pour mettre en correspondance les points de contour détectés sur les droites épipolaires d'une séquence d'images stéréoscopiques réelles, nous décrirons le processus génétique mis en œuvre pour appairer ces primitives.

3.2 Algorithmes génétiques

Depuis que Charles Darwin a formulé l'idée simple d'une évolution naturelle des espèces basée sur un mécanisme de sélection et de reproduction, la théorie de l'évolution biologique est devenue un centre d'intérêt pour les scientifiques. On sait que l'évolution d'une espèce a lieu lorsque les chromosomes, qui contiennent le code génétique de sa structure biologique, se trouvent modifiés. Bien que toutes les caractéristiques des évolutions naturelles ne soient pas connues, les spécialistes considèrent qu'il faut admettre au moins les règles suivantes :

- L'évolution n'agit pas directement sur les êtres vivants, elle opère en réalité sur les chromosomes contenus dans leurs cellules.
- L'évolution est régie par deux phénomènes : la sélection et la reproduction. La sélection naturelle construit un lien entre les chromosomes et les individus. Elle permet d'aboutir à un meilleur taux de reproduction des individus qui ont de "bons" chromosomes.
- La reproduction est le moteur de l'évolution. Les modifications apparaissant lors de la reproduction peuvent produire de nouvelles générations dont les caractéristiques sont différentes.
- L'ensemble des chromosomes avec leur codage comprennent toutes les informations caractéristiques d'un individu.

La plupart des espèces biologiques se reproduisent par une combinaison d'une sélection naturelle et d'une reproduction sexuée. La sélection décide quel individu peut vivre et se reproduire alors que la reproduction sexuée assure le "mélange" du matériel génétique. Les espèces qui utilisent un mode de reproduction non sexué évoluent en général beaucoup moins vite, car la modification du matériel génétique n'intervient que lors des mutations.

Ces principales caractéristiques de l'évolution naturelle ont attiré John Holland et ses collègues de l'université du Michigan lors de recherches sur l'apprentissage des machines dans les années 60. Holland a trouvé que l'apprentissage se réalise non seulement par l'adaptation

d'un individu, mais également par l'évolution d'un groupe lors de générations successives. En 1975, il a publié son ouvrage "Adaptation in Natural and Artificial Systems" dans lequel il a proposé les premières versions d'un algorithme génétique qui permet un apprentissage basé sur les règles de la sélection naturelle.

3.2.1 Méthodes de sélection

Dans la littérature, on trouve plusieurs méthodes de sélection [Mic94]. Dans ce qui suit, nous allons en décrire quatre.

- La sélection équiprobable.

Cette méthode consiste à choisir les couples "parents" de manière équiprobable parmi les chromosomes de la population. La fonction d'évaluation n'intervient pas dans le choix de ces chromosomes.

- La sélection déterministe.

Cette méthode trie les chromosomes selon les valeurs prises par la fonction objectif. Seule une partie des meilleurs chromosomes est sélectionnée pour la phase de reproduction. Cette méthode pourrait induire une convergence prématurée de l'algorithme car la pression de sélection est alors très forte. Il est en effet nécessaire de maintenir une diversité génétique suffisante dans la population, celle-ci constituant un réservoir de gènes pouvant être utile par la suite. Tout chromosome peut transmettre à sa descendance des gènes qui, une fois combinés avec d'autres, peuvent se révéler intéressants.

- La sélection par tournoi.

Deux chromosomes sont choisis au hasard. Le plus adapté accède à la phase de reproduction. Il est tout à fait possible que certains chromosomes participent à plusieurs tournois. S'ils gagnent plusieurs fois, ils auront donc le droit d'être copiés plusieurs fois, ce qui favorisera la pérennité de leurs gènes.

- La sélection stochastique ou non-déterministe, encore dite sélection par la méthode de la roulette de casino.

Selon cette méthode, chaque chromosome occupe un secteur d'une roulette dont l'angle est proportionnel à sa probabilité de sélection. Celle-ci est d'autant plus grande que la valeur de la fonction objectif associée au chromosome est grande. Comme dans la sélection par tournoi, certains chromosomes peuvent être sélectionnés plusieurs fois pour accéder

à la phase de reproduction.

Notons qu'il est possible d'associer la procédure d'élitisme aux deux dernières techniques de sélection par tournoi et par roulette de casino. Cette procédure est appliquée de la façon suivante : si par hasard le chromosome le mieux adapté, qui représente la meilleure solution, n'a pas été sélectionné, alors il remplace un chromosome choisi aléatoirement avant d'accéder à la phase suivante, à savoir la reproduction.

Une fois la phase de sélection des chromosomes terminée, la phase de reproduction peut commencer. Pour ce faire, les chromosomes sont tout d'abord regroupés aléatoirement deux par deux. Ensuite, des opérations de croisement et de mutation leurs sont appliquées avec une probabilité à définir.

3.2.2 Opérateur de croisement

Le croisement est une opération sexuée qui assure le brassage et la recombinaison des gènes parentaux, pour former des descendants aux potentialités nouvelles. En d'autres termes, cet opérateur favorise l'exploration de l'espace de recherche grâce aux chromosomes enfants conçus à partir des gènes provenant de chromosomes parents.

Avant de croiser deux chromosomes, une valeur est tirée au sort entre 0 et 1. Si cette valeur est inférieure ou égale à la probabilité de croisement, les deux chromosomes sont désignés comme reproducteurs. Pour les croiser, nous disposons de plusieurs méthodes. Nous en présentons trois :

- Le croisement en un point.

Le croisement en un point de deux chromosomes est réalisé de la façon suivante. On choisit aléatoirement un point de croisement situé à la position p comprise entre 1 et $l_c - 1$, où l_c est la longueur du chromosome. Afin de créer les deux nouveaux chromosomes, on conserve intacts les gènes compris entre les positions 1 et p , et on échange les gènes compris entre les positions $p + 1$ et l_c . Ce procédé est illustré sur la figure 3.1

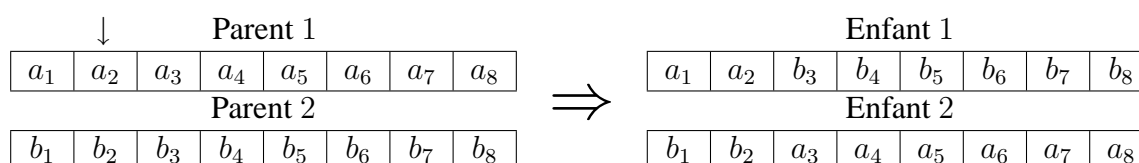


FIG. 3.1 : Croisement en position 2 de deux chromosomes parents de longueur 8.

- Le croisement en deux points.

On choisit au hasard deux points de croisement. Selon cette méthode, on échange les gènes qui sont situés entre les deux points de croisement tirés au hasard.

- Le croisement uniforme.

Le croisement uniforme est radicalement différent du croisement en un point. En effet, dans cette méthode, chaque gène du chromosome enfant est créé en copiant le gène correspondant de l'un des deux parents, selon un masque généré aléatoirement. Là où il y a 1 dans le masque de croisement, le gène est copié du premier parent et là où il y a 0 dans le masque, le gène est copié du deuxième (cf. figure 3.2). Le processus est répété en échangeant les parents pour produire le deuxième enfant. Les enfants contiennent donc un mélange des gènes de chacun de leurs parents.

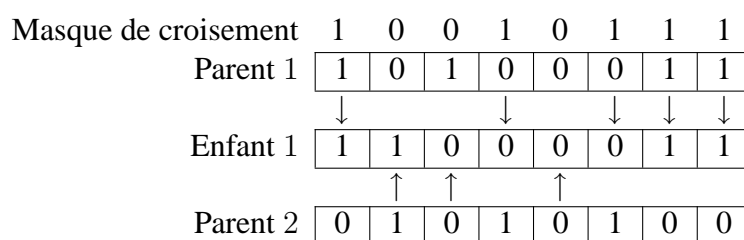


FIG. 3.2 : Croisement uniforme.

Notons que d'autres formes de croisement existent, du croisement en k points jusqu'au cas limite du croisement uniforme.

3.2.3 Opérateur de mutation

La mutation classique consiste à modifier la valeur d'un gène tiré d'une manière aléatoire dans un chromosome. Dans le cas d'un chromosome binaire, cette modification n'est rien d'autre qu'une substitution de la valeur du gène tiré aléatoirement par une autre. Si cette valeur est à 1, alors elle est remplacée par 0 et vice-versa (cf. figure 3.3). Les mutations empêchent l'évolution de se figer. Elles permettent d'assurer une recherche aussi bien globale que locale, selon le poids et le nombre des gènes mutés. Toutefois, pour ne pas trop perturber l'évolution globale de la population, la probabilité d'apparition d'une mutation doit rester très faible.

D'autre part, une population trop petite peut s'homogénéiser à cause des erreurs stochastiques : les gènes favorisés par le hasard peuvent se répandre au détriment des autres. Cela signifie que l'on risque alors d'aboutir à des solutions qui ne seront pas forcément optimales.

Les mutations permettent de contrebalancer cet effet en introduisant constamment de nouveaux gènes dans la population.

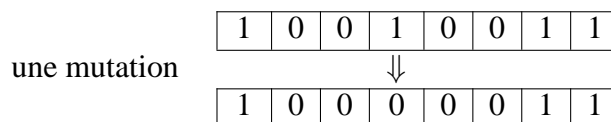


FIG. 3.3 : Représentation schématique d'une mutation dans un chromosome binaire.

3.2.4 Convergence des AGs

La convergence prématurée est l'un des problèmes majeurs rencontrés dans les algorithmes itératifs. Dans le cas des algorithmes génétiques, on appelle convergence prématurée le fait d'aboutir à une population stable de chromosomes qui ont convergé vers un optimum local de la fonction d'évaluation plutôt que vers un optimum global. Cela se produit lorsque l'évolution de la population a été trop rapide, ce qui n'a pas permis d'explorer suffisamment l'espace des solutions.

Ce problème est fortement lié au choix des paramètres régissant le fonctionnement de l'algorithme génétique, c'est-à-dire la taille de la population et les probabilités de croisement et de mutation. En effet, l'opération de sélection assure la convergence vers une bonne solution alors que les opérations de croisement et de mutation permettent d'explorer l'espace de solutions. Si on favorise l'opération de sélection, on peut aboutir à une convergence prématurée. Par contre, si on favorise les opérations de croisement et de mutation, l'algorithme peut ne pas converger.

Plusieurs méthodes permettant de choisir les paramètres ont été décrites dans la littérature. Certaines techniques consistent à fixer les paramètres qui restent constants pendant toute la durée du processus d'optimisation [Jon75, Gre86, SCED89]. Pour améliorer la convergence, plusieurs auteurs ont proposé d'ajuster dynamiquement les valeurs des paramètres durant l'évolution de la population [HM91, Bäck92, AMM94, SP94, XV94, GAF95, HLV97, Bäck96, HL96, HYFS96, WWSJ96, Wu98].

3.2.5 Conclusion

Les AGs s'inspirent du mode d'évolution naturel. Ils imitent le processus d'évolution d'une population d'individus ou de chromosomes. L'un des problèmes majeurs dans l'application des

AGs réside dans le codage de la solution du problème sous forme de chromosomes. Chaque chromosome est composé de plusieurs gènes dont le indique la taille de celui-ci.

Une fois que le codage du problème est établi, l'ensemble des moyens nécessaires à la mise en œuvre d'un AG nécessite :

- un moyen de créer la population initiale,
- une fonction d'évaluation pour mesurer l'adaptation de chaque chromosome par rapport à un objectif fixé,
- un mode de sélection des chromosomes à reproduire,
- des opérateurs génétiques adaptés au problème,
- des valeurs pour les paramètres qu'utilise l'algorithme (taille de la population, probabilités d'appliquer les différents opérateurs, etc),
- un critère d'arrêt.

Dans la suite de ce chapitre, nous décrirons le codage proposé pour l'appariement d'images-ligne épipolaires ainsi que les différents mécanismes nécessaires à la mise en œuvre d'un AG pour résoudre le problème de mise en correspondance d'indices visuels entre les images-ligne épipolaires droite et gauche.

3.3 Codage binaire du problème de mise en correspondance

Dans ce paragraphe, nous allons décrire le codage génétique mis en œuvre afin de mettre en correspondance les points de contour extraits des couples d'images-ligne épipolaires gauche et droite. Soit L_G (resp. L_D) la liste des points de contour extraits de l'image-ligne gauche (resp. droite). Chaque élément de ces listes est un triplet (x, s_x, d_x) où x désigne la position du point de contour dans l'image-ligne épipolaire correspondant, s_x représente le signe de la dérivée en ce point qui prend les valeurs $+1$ ou -1 ¹ et d_x est l'amplitude de la dérivée en ce point. Chacune de ces deux listes est ordonnée de manière croissante selon les positions des points de contour extraits de l'image-ligne épipolaire correspondant. Le point de contour ayant la position la plus proche du début de la ligne est placé en tête de la liste et ainsi de suite.

Soit G (resp. D) la liste des indices des points de contour de la liste L_G (resp. L_D). Soit N_G (resp. N_D) le nombre de points de contour extraits de l'image-ligne gauche (resp. droite).

Le principe de base du codage est de représenter le problème en termes d'états qui sont

¹Si $s_x = -1$ alors il s'agit d'un front descendant, sinon il s'agit d'un front montant

représentés par des variables binaires. Chacun de ces états représente une possibilité de mise en correspondance entre un point de contour de l'image-ligne épipolaire gauche et un point de contour de l'image-ligne épipolaire droite. Nous désignons par E_{gd} la variable d'état qui représente la possibilité d'appariement d'un point de contour g de l'image-ligne épipolaire gauche avec un point de contour d de l'image-ligne épipolaire droite (cf. figure 3.4) :

$$E_{gd} = \begin{cases} 1 & \text{si l'appariement entre } g \text{ et } d \text{ est possible} \\ 0 & \text{sinon} \end{cases}$$

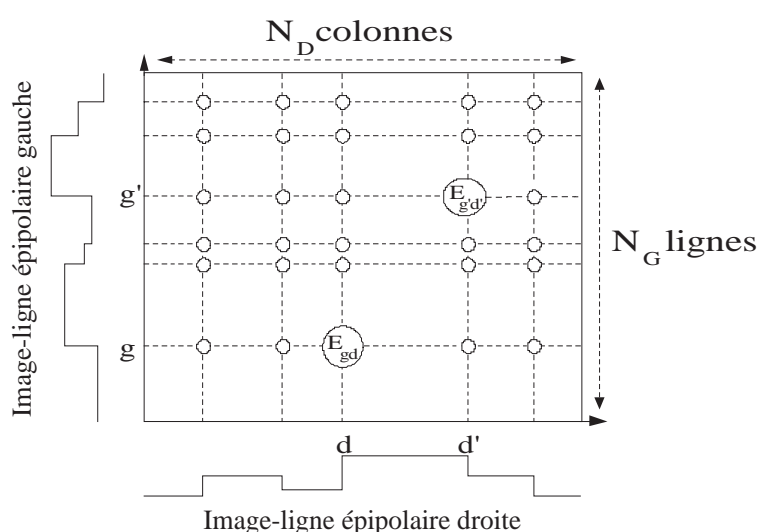


FIG. 3.4 : Schéma de mise en correspondance des points de contour d'un couple d'images-ligne épipolaires gauche et droite.

Ainsi, le codage du problème de la mise en correspondance pour un couple d'images-ligne épipolaires nécessite un système de $N_G \times N_D$ variables d'état. Pour permettre une bonne compréhension, nous nous aidons du tableau de la figure 3.5 que nous appelons tableau de mise en correspondance. Dans ce tableau, nous supposons que le nombre de points de contour extraits de l'image-ligne épipolaire gauche est de 5 ($N_G = 5$). Ainsi la liste G est composée des éléments $\{1, 2, 3, 4, 5\}$. Nous supposons aussi que le nombre de points de contour extraits de la ligne épipolaire droite est de 6 ($N_D = 6$). Ainsi la liste D est composée des éléments $\{1, 2, 3, 4, 5, 6\}$. Un élément de ce tableau, comme E_{34} par exemple, désigne la possibilité d'appariement du point de contour d'indice 3 de la liste G avec le point de contour d'indice 4 de la liste D . Si $E_{34} = 1$, cela signifie que les points sont appariés. Sinon ils ne le sont pas.

À partir du tableau de mise en correspondance, nous pouvons facilement construire un chromosome sous forme d'une chaîne binaire en parcourant les éléments du tableau de bas en haut et

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

G

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

FIG. 3.5 : Tableau de mise en correspondance des listes des indices des points de contour G et D .

de gauche à droite. Ainsi, les éléments E_{gd} seront les gènes qui forment le chromosome binaire (cf. figure 3.6). Notons que ce passage d'une représentation matricielle à une représentation chromosomique ne fait perdre aucune information. En effet, la position p d'un gène dans le chromosome permet de trouver les indices g et d des points de contour candidats à l'appariement grâce aux formules suivantes :

$$g = 1 + Ent(p - 1, N_D),$$

$$d = 1 + Modulo(p - 1, N_D),$$

où Ent désigne l'opérateur de la division entière et $Modulo$ désigne celui du reste de la division. Grâce à ce codage, les solutions pour la mise en correspondance des points de contour des lignes épipolaires d'images stéréoscopiques sont représentées sous forme de chromosomes binaires. A présent, il faut définir une fonction d'évaluation, ou fonction objectif, à optimiser qui permet d'évaluer la qualité de chaque solution par rapport au problème traité. Cette fonction doit être construite de telle manière que les "bons" appariements correspondent à ses extrema, minima ou maxima.

E ₁₁	E ₁₂	E ₁₃	E ₁₄	E ₁₅	E ₁₆	E ₅₁	E ₅₂	E ₅₃	E ₅₄	E ₅₅	E ₅₆
-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------

FIG. 3.6 : Chromosome formé à partir du tableau de mise en correspondance.

L'approche que nous avons adoptée est basée sur une démarche de recherche globale [Rui97].

Le problème de la mise en correspondance est considéré comme un problème de satisfaction de contraintes. La solution optimale au problème est celle pour laquelle l'ensemble des appariements respecte au mieux ces contraintes. Pour quantifier cette optimalité, il suffit de construire une fonction d'évaluation H qui est une application de $\{0, 1\}^{N_G \times N_D^2}$ dans R . Représentant les contraintes du problème, celle-ci associe à chaque possibilité d'appariement un nombre réel qui représente la qualité de la solution en question ou le degré de compatibilité des appariements vis-à-vis des contraintes imposées. Il s'agit alors de chercher la solution qui correspond à un extremum de cette fonction d'évaluation. Les différentes contraintes sont détaillées dans le paragraphe suivant.

3.4 Contraintes de mise en correspondance

3.4.1 Contraintes locales

Les contraintes locales sont exprimées en fonction des attributs des primitives gauches et droites, prises deux à deux. Elles sont dites locales car elles sont appliquées à chaque couple de primitives candidates à l'appariement, sans faire intervenir les autres couples. Nous utilisons deux contraintes locales : la contrainte de position et la contrainte de signe de la dérivée.

3.4.1.1 Contrainte de position

La géométrie d'un stéréoscope impose une condition sur les positions relatives du point de contour g issu de la ligne épipolaire gauche x_g et du point de contour d issu de la ligne épipolaire droite x_d (cf. figure 3.7). Les points du couple (g, d) ne pourront être appariés que si leurs abscisses respectives vérifient la condition : $x_g > x_d$. Ainsi, les points de tout couple ne respectant pas cette conditions sont considérés comme incompatibles.

3.4.1.2 Contrainte de signe de la dérivée

Le long d'une ligne épipolaire, une dérivée positive correspond à un front montant, c'est-à-dire à une variation du niveau de gris évoluant du sombre vers le clair quand on explore la ligne selon les abscisses croissantes. Par contre, une dérivée négative indique un front descendant, c'est-à-dire une variation du niveau de gris d'une zone claire vers une zone sombre.

²L'ensemble des $(N_G \times N_D)$ -uplets dont les éléments prennent les valeurs 0 ou 1.

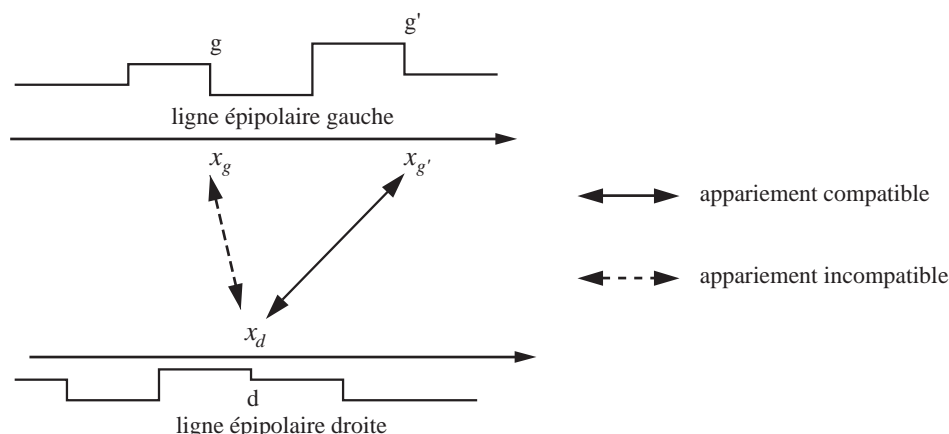


FIG. 3.7 : Contrainte de position.

La contrainte de signe du gradient stipule qu'un front montant ne peut être mis en correspondance avec un front descendant (cf. figure 3.8). Ainsi, les points de contour de tout couple ne respectant pas cette contrainte sont considérés comme incompatibles.

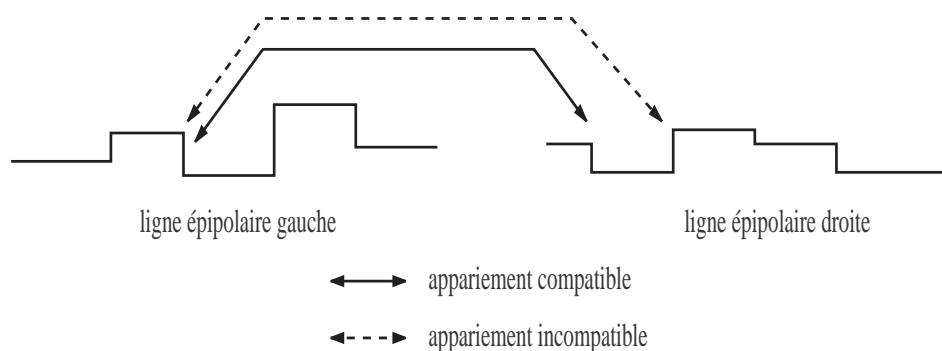


FIG. 3.8 : Contrainte de signe de la dérivée.

3.4.2 Contraintes globales

Contrairement aux contraintes locales, les contraintes globales impliquent tous les couples de primitives candidates à l'appariements pour réduire les ambiguïtés de mise en correspondance. Nous utilisons trois contraintes globales : la contrainte d'unicité, la contrainte d'ordre et la contrainte de continuité de la disparité.

3.4.2.1 Contrainte d'unicité

La contrainte d'unicité constitue l'un des aspects fondamentaux de la stéréovision. Elle stipule que toute primitive de la ligne épipolaire gauche a au plus un correspondant dans la

ligne épipolaire droite, et réciproquement.

3.4.2.2 Contrainte d'ordre

La contrainte d'ordre suppose que l'ordre des points de contour est le même dans chaque ligne épipolaire. Soient deux couples (g, d) et (g', d') de points de contour homologues (cf. figure 3.9). La compatibilité de ces deux couples, au sens de la contrainte d'ordre, est vérifiée si la position, dans la ligne épipolaire gauche, du point de contour g par rapport au point de contour g' est la même que celle du point d par rapport au point d' , dans la ligne épipolaire droite. Dans certaines scènes, cette contrainte peut ne pas être respectée, comme illustré sur la figure 3.10.

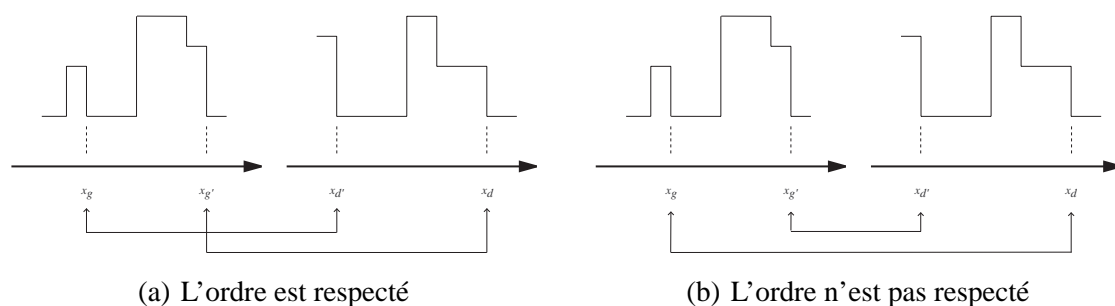


FIG. 3.9 : La contrainte d'ordre

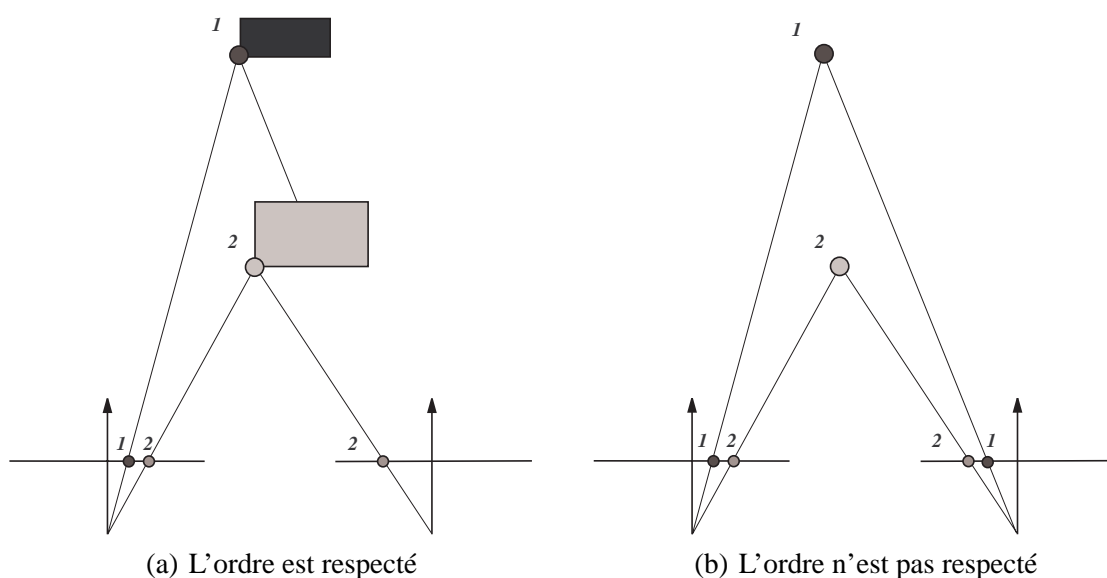


FIG. 3.10 : Cas particulier où la contrainte d'ordre n'est pas respectée

3.4.2.3 Contrainte de continuité de la disparité

La contrainte de continuité de la disparité s'appuie sur l'hypothèse de continuité des surfaces par morceaux de telle sorte qu'un même objet ne peut présenter de variations brutales de profondeur par rapport au stéréoscope. La profondeur étant directement liée à la disparité, il en résulte que les disparités doivent varier continûment le long des surfaces constituant les objets. La contrainte de continuité de la disparité permet ainsi de vérifier la cohérence des appariements dans un contexte global.

3.4.3 Conclusion

Notre procédure de mise en correspondance utilise les cinq contraintes stéréoscopiques que nous venons de présenter :

1. La contrainte de position ;
2. La contrainte de signe de la dérivée ;
3. La contrainte d'unicité ;
4. La contrainte d'ordre ;
5. La contrainte de continuité de la disparité.

Nous allons montrer comment la prise en compte des deux premières contraintes permet de ne coder que des chromosomes qui représentent des solutions a priori compatibles avec les contraintes locales de position et de signe de la dérivée.

3.5 Amélioration du codage

Comme nous l'avons déjà montré, pour résoudre le problème de mise en correspondance en utilisant un algorithme génétique, nous devons trouver un codage approprié au problème traité. Dans le codage proposé au paragraphe 3.3, une solution du problème de mise en correspondance peut-être représentée sous forme d'un tableau que nous avons appelé le tableau de mise en correspondance. La dimension de celui-ci est de $(N_G \times N_D)$ où N_G et N_D sont les nombres totaux de points de contour détectés, respectivement dans les images-ligne épipolaires gauche et droite (cf. figure 3.4). Pour rendre l'exploration de l'espace de recherche des solutions plus efficace, nous nous servirons des contraintes locales pour produire directement des chromosomes binaires représentant des appariements compatibles vis-à-vis de celles-ci (cf. figure 3.11). Ainsi,

les gènes incompatibles avec ces contraintes locales ne seront pas introduits dans la chaîne binaire représentant le code génétique d'une solution de mise en correspondance. Pour mieux illustrer cette amélioration apportée au codage binaire, nous nous servons d'un exemple.

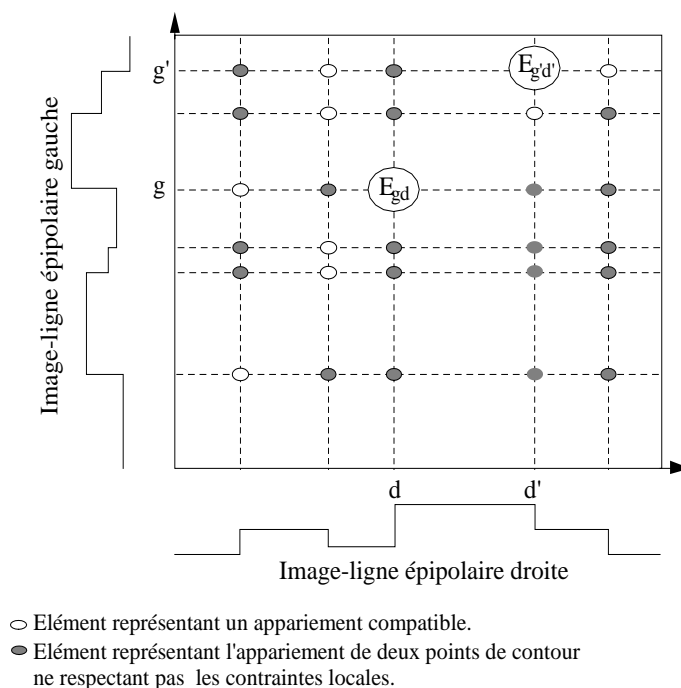


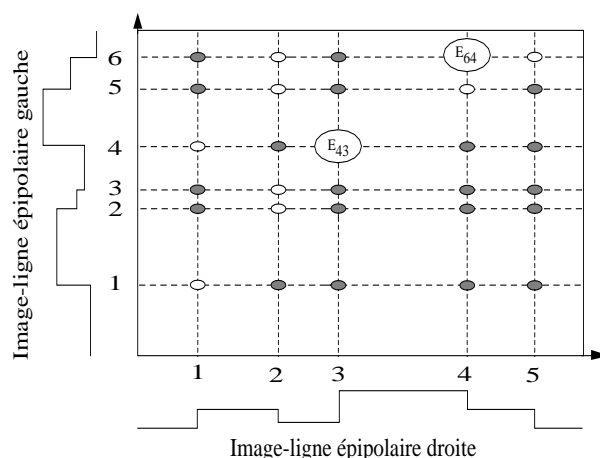
FIG. 3.11 : Éléments parmi lesquels un chromosome est formé.

Supposons que $N_G = 6$ et que $N_D = 5$ (cf. figure 3.12(a)). Pour former le chromosome binaire compacte, nous parcourons le schéma de mise en correspondance de gauche à droite et de bas en haut en évitant les éléments représentant l'appariement de deux points de contour ne respectant pas les contraintes locales. De ce fait, chaque gène doit être composé de trois éléments : l'indice g du point de contour dans la liste gauche G , l'indice d du point de contour dans la liste droite D ainsi que la valeur binaire de E_{gd} qui indique si ces deux points sont appariés ($E_{gd} = 1$) ou non ($E_{gd} = 0$). Ainsi, un chromosome a la forme illustrée dans la figure 3.12(b).

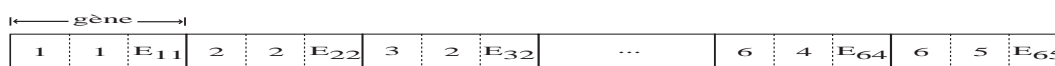
La prochaine étape consiste à construire une fonction d'évaluation de telle sorte que ses minima correspondent à des appariements corrects vis-à-vis les cinq contraintes imposées.

3.6 Fonction d'évaluation

La fonction d'évaluation du problème de mise en correspondance est définie uniquement à partir des termes associés à chacune des contraintes globales. En effet, les contraintes locales de



(a) Schéma de mise en correspondance de couple d'images-ligne épipolaires avec $N_G = 6$ et $N_D = 5$.



(b) Chromosome formé uniquement des gènes respectant les contraintes locales.

FIG. 3.12 : Exemple d'un chromosome compacte pour la mise en correspondance d'un couple d'images-ligne épipolaires avec $N_G = 6$ et $N_D = 5$.

position et de signe de la dérivée ont été prises en compte lors de la phase d'initialisation de la population de chromosomes compacts. Dans les paragraphes suivants, nous décrivons les trois termes dont la fonction d'évaluation est composée.

3.6.1 Fonction d'évaluation représentant la contrainte d'unicité

Reprenons la représentation décrite au paragraphe 3.3. Une ligne (respectivement une colonne) du tableau de mise en correspondance représente l'ensemble des appariements possibles d'un point de contour de l'image-ligne épipolaire gauche (respectivement droite). Pour satisfaire la contrainte d'unicité, il faut d'abord que chaque ligne du tableau de mise en correspondance mette en évidence un seul appariement, c'est-à-dire qu'il doit y avoir une seule case à 1 et les autres cases à 0. Le même raisonnement s'applique également pour les colonnes. La figure 3.13(a) représente un exemple d'appariements respectant la contrainte d'unicité. Cette dernière n'est pas respectée pour l'exemple de la figure 3.13(b) car la présence de deux 1 sur la ligne 4 du tableau de mise en correspondance indique que le point de contour 4 de l'image-ligne épipolaire gauche est apparié aux deux points de contour 3 et 5 de l'image-ligne épipolaire droite.

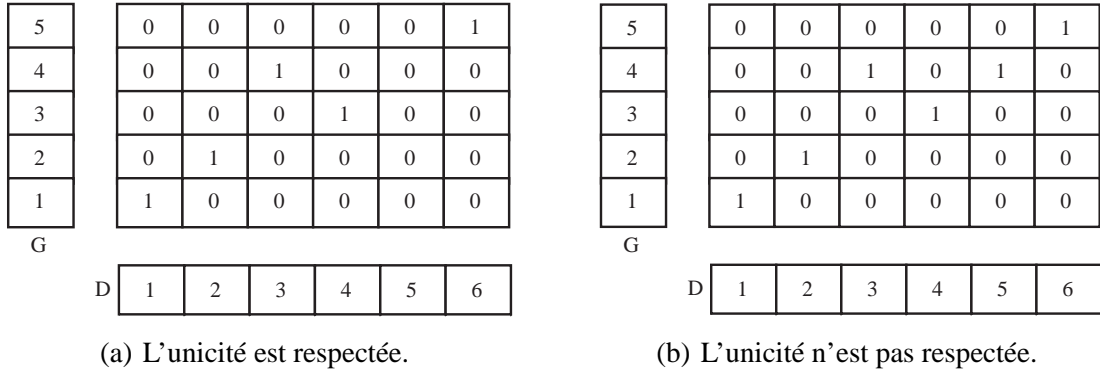


FIG. 3.13 : Codage de la contrainte d'unicité

La fonction d'évaluation représentant la contrainte d'unicité doit être construite de telle façon que les possibilités d'appariement respectant cette contrainte soient favorisées lorsqu'elle atteint son minimum. Cette fonction peut être exprimée sous la forme :

$$H_{uni} = \left[\sum_{g \in G} \left(1 - \sum_{d \in D / (g,d) \in \Omega} E_{gd} \right)^2 + \sum_{d \in D} \left(1 - \sum_{g \in G / (g,d) \in \Omega} E_{gd} \right)^2 \right]. \quad (3.1)$$

où Ω représente l'ensemble des appariements possibles après prise en compte des contraintes de position et de signe de la dérivée :

$$\Omega = \{(g, d) \in G \times D / g \text{ et } d \text{ vérifient les contraintes de position et de signe de la dérivée}\}$$

Le premier terme de cette fonction sert à défavoriser les situations où il y a plus d'un 1 sur chaque image-ligne du tableau de mise en correspondance. Il contribue à n'autoriser qu'un seul appariement pour chaque point de contour de l'image-ligne gauche. Le second terme sert à défavoriser les situations où il y a plus d'un 1 sur chaque colonne. Il contribue à n'autoriser qu'un seul appariement pour chaque point de contour de l'image-ligne droite.

3.6.2 Fonction d'évaluation représentant la contrainte d'ordre

Considérons deux couples (g, d) et (g', d') de points de contour homologues. La compatibilité des appariements de ces deux couples n'est vérifiée que si l'ordre des points de contour g et g' dans l'image-ligne épipolaire gauche est le même que celui des points de contour d et d' dans l'image-ligne épipolaire droite.

La figure 3.14(a) montre un exemple où la contrainte d'ordre est respectée. En suivant l'ordre croissant des points de contour dans chaque image-ligne, les appariements sont disposés

de telle sorte qu'ils constituent un chemin ascendant. Un exemple de possibilité d'appariement où la contrainte d'ordre est mise en défaut est illustré sur la figure 3.14(b).

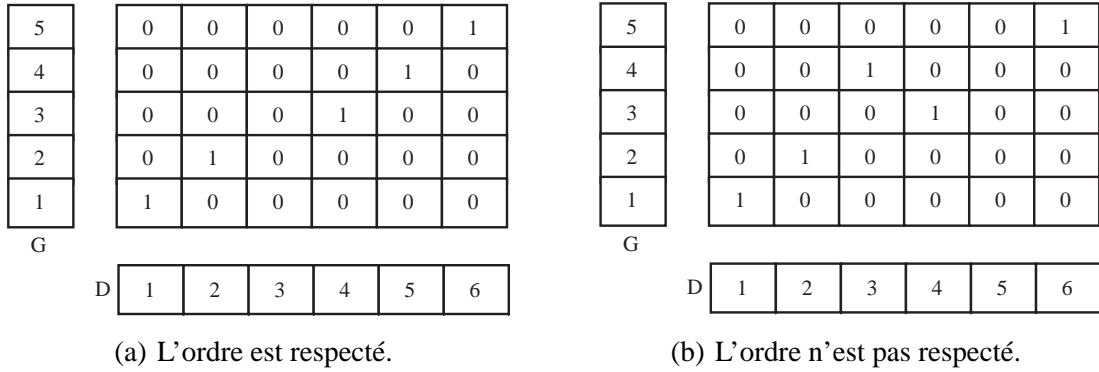


FIG. 3.14 : Codage de la contrainte d'ordre

Pour construire une fonction d'évaluation représentant la contrainte d'ordre, il faut que les possibilités de mise en correspondance vérifiant cette contrainte soient favorisées lorsqu'elle atteint son minimum. Cette fonction peut s'écrire sous la forme :

$$H_{ord} = \sum_{(g,d) \in \Omega} \sum_{\substack{(g',d') \in \Omega \\ g' \neq g \text{ et } d' \neq d}} O_{gdg'd'} E_{gd} E_{g'd'}. \quad (3.2)$$

Le coefficient $O_{gdg'd'}$, qui indique si la contrainte d'ordre est respectée pour les deux appariements (g, d) et (g', d') , est calculé comme suit :

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

où $x_g, x_d, x_{g'}$ et $x_{d'}$ représentent respectivement les abscisses des points de contour g, d, g' et d' dans les images-ligne stéréoscopiques. $S(a)$ est la fonction signe telle que :

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

3.6.3 Fonction d'évaluation représentant la contrainte de continuité

Nous avons vu dans le paragraphe 3.4.2.3 que la contrainte de continuité est fondée sur l'hypothèse de continuité des surfaces. Ainsi, les couples de primitives appartenant à un même objet doivent avoir des disparités voisines. La fonction d'évaluation associée à cette contrainte favorise les couples de primitives candidats à l'appariement dont la différence de disparité avec

les autres couples est inférieure à un certain seuil. Cette fonction peut être exprimée sous la forme :

$$H_{cont} = - \sum_{(g,d) \in \Omega} \sum_{\substack{(g',d') \in \Omega \\ g' \neq g \text{ et } d' \neq d}} S_{gdg'd'} E_{gd} E_{g'd'} . \quad (3.3)$$

Le coefficient $S_{gdg'd'}$ représente une mesure de compatibilité entre les deux paires de points de contour (g, d) et (g', d') . Il est exprimé en utilisant la fonction non linéaire suivante [NC92] :

$$S_{gdg'd'} = S(X) = \frac{2}{1 + e^{(X-\theta)}} - 1, \quad (3.4)$$

où X est la valeur absolue de la différence des disparités entre les deux paires (g, d) et (g', d') :

$$X = | (x_g - x_d) - (x'_g - x'_d) | .$$

Le paramètre θ permet une certaine tolérance afin de prendre en compte les effets du bruit et les distorsions (cf. figure 3.15). Il est ajusté de telle sorte qu'une compatibilité maximale, c'est-à-dire proche de 1, soit obtenue pour les bons appariements qui correspondent à des valeurs de X voisines de zéro, et qu'une compatibilité minimale, c'est à dire proche de -1 , soit obtenue pour les mauvais appariements qui correspondent à des valeurs élevées de X . Les tests expérimentaux ont montré que $\theta = 20$ est une valeur satisfaisante pour ce paramètre afin de prendre en compte les effets du bruit et les distorsions [Rui97].

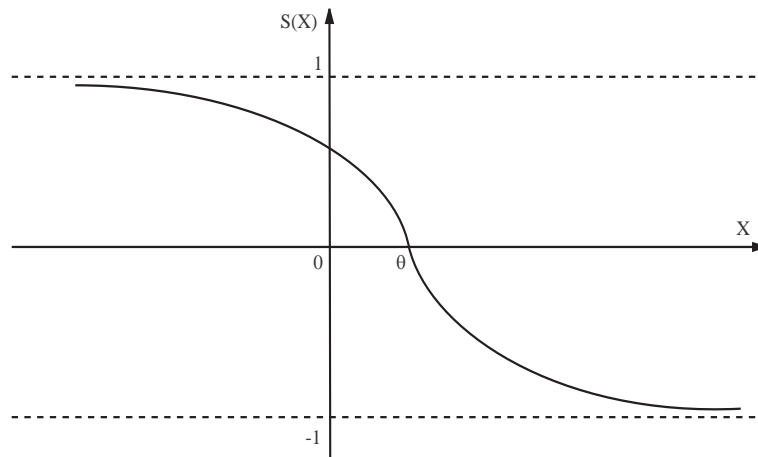


FIG. 3.15 : Graphe de la fonction non linéaire $S(X)$.

3.6.4 Fonction d'évaluation totale

En combinant les trois termes des équations 3.1, 3.2 et 3.3 et en ajoutant un terme favorisant les chromosomes codant un nombre maximal d'appariements, la fonction d'évaluation totale

H_{tot} associée au problème de mise en correspondance prend la forme suivante :

$$\begin{aligned}
H_{tot} = & K_{uni} \left[\sum_{g \in G} \left(1 - \sum_{d \in D / (g,d) \in \Omega} E_{gd} \right)^2 + \sum_{d \in D} \left(1 - \sum_{g \in G / (g,d) \in \Omega} E_{gd} \right)^2 \right] + \\
& K_{ord} \sum_{(g,d) \in \Omega} \sum_{\substack{(g',d') \in \Omega \\ g' \neq g \text{ et } d' \neq d}} O_{gdg'd'} E_{gd} E_{g'd'} - \\
& K_{cont} \sum_{(g,d) \in \Omega} \sum_{\substack{(g',d') \in \Omega \\ g' \neq g \text{ et } d' \neq d}} S_{gdg'd'} E_{gd} E_{g'd'} + \\
& K_m \left(N_{min} - \sum_{g \in G} \sum_{d \in D / (g,d) \in \Omega} E_{gd} \right)^2, \tag{3.5}
\end{aligned}$$

où $N_{min} = \min(N_G, N_D)$ et où K_{uni} , K_{ord} , K_{cont} , K_m sont des coefficients positifs permettant de pondérer les contributions des différents termes.

3.7 Mise en œuvre de l'algorithme génétique

Le codage mis en œuvre et la fonction d'évaluation construite nous ont permis de ramener le problème de mise en correspondance à un problème d'optimisation.

La fonction d'évaluation est construite de telle sorte que ses minima correspondent aux "bons" appariements. L'algorithme génétique que nous décrivons dans ce chapitre est destiné à minimiser cette fonction d'évaluation. A l'issue du processus d'optimisation, la solution de notre problème correspond au minimum trouvé.

Les étapes successives de l'algorithme mis en œuvre sont présentées sur la figure 3.16.

Étape 1 : Comme dans tous les AGs, la première étape est une étape d'initialisation. Elle consiste à créer aléatoirement une population de chromosomes de taille $Taille_{pop}$ à partir de laquelle l'algorithme démarre. Nous rappelons que tous les chromosomes créés lors de cette phase respectent les contraintes stéréoscopiques locales de position et de signe de la dérivée.

Étape 2 : Comme le montre l'organigramme de la figure 3.16, la deuxième étape est destinée à l'évaluation des chromosomes de la population initiale grâce à la fonction d'évaluation définie par l'équation 3.5.

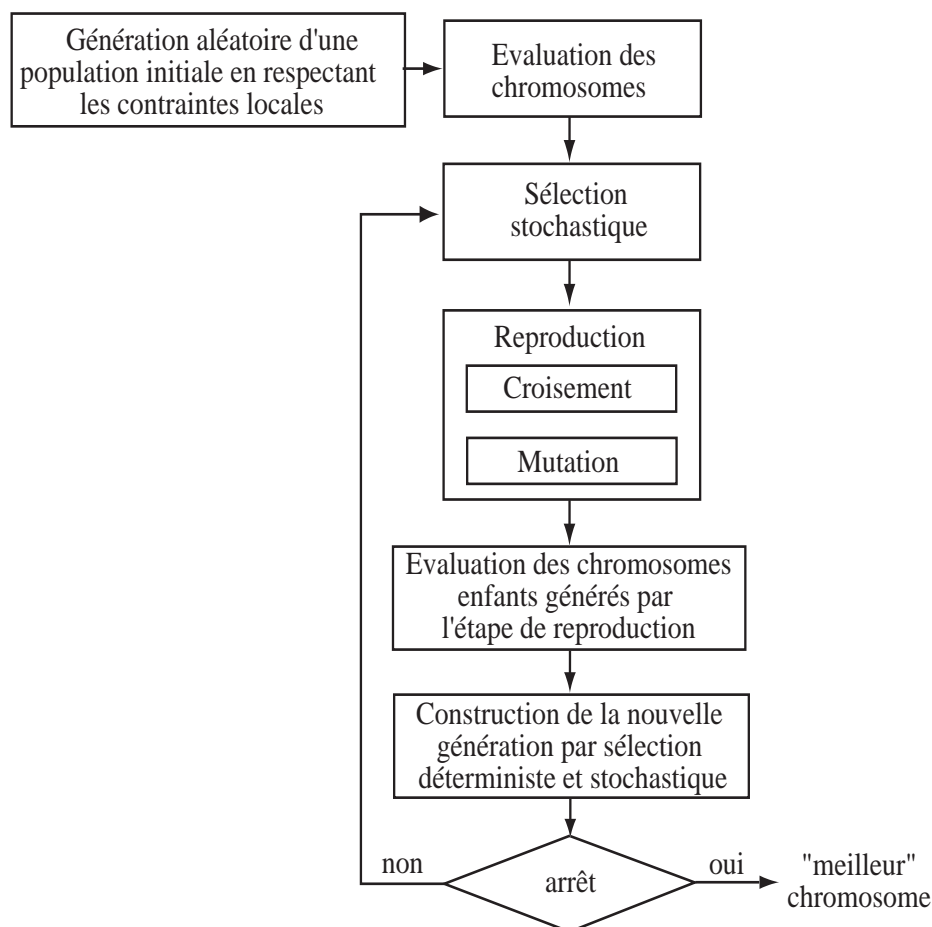


FIG. 3.16 : Schéma de l'algorithme génétique.

Étape 3 : Dans la troisième étape, nous procédons à la phase de sélection des chromosomes qui pourraient jouer le rôle des parents s'ils sont admis à la reproduction. Cette sélection stochastique est basée sur le principe de la roulette de casino.

Étape 4 : Les chromosomes sélectionnés à l'étape précédente sont ensuite regroupés par paires pour effectuer leur croisement. Il s'agit, pour chaque couple de chromosomes sélectionnés, de tirer une valeur aléatoire comprise entre 0 et 1. Si cette valeur est égale ou supérieure à la probabilité de croisement P_{crois} les deux chromosomes sont croisés. Sinon les deux chromosomes restent intacts. Parmi les différentes méthodes de croisement qui existent (cf. § 3.2.2), l'opérateur de croisement en un point est retenu car il est simple à mettre en œuvre et permet de générer des chromosomes enfants qui respectent les contraintes locales comme leurs parents. Dans le cadre de cette thèse nous n'avons pas testé l'influence du choix de l'opérateur de croisement sur le résultat de la reconstruction 3D des images stéréoscopiques traitées.

Étape 5 : Dans cette étape, on procède à la mutation des chromosomes issus de l'étape 4 avec une probabilité P_{mut} qui est la probabilité de muter un gène dans un chromosome. En effet, la mutation d'un chromosome est réalisée de la manière suivante : il s'agit de parcourir tous les gènes qui constituent un chromosome et de tirer aléatoirement, pour chacun de ces gènes, une valeur v comprise entre 0 et 1 puis de substituer, si $v \leq P_{mut}$, la valeur binaire de l'élément E_{gd} du gène par son complément ³. Nous attirons l'attention du lecteur sur le fait que la mutation d'un gène ne concerne que son élément E_{gd} . Les deux autres éléments qui forment celui-ci, à savoir l'indice g du point de contour dans la liste gauche G et l'indice d du point de contour dans la liste droite D ne subissent aucune mutation.

Étape 6 : Dans la sixième étape, les chromosomes issus de la phase de reproduction, qui constituent la population d'enfants, sont évalués grâce au formule 3.5 pour déterminer leur adaptation à l'objectif recherché. Nous signalons que dans notre approche les chromosomes "enfants" ne remplacent pas systématiquement les chromosomes "parents" comme c'est le cas dans un processus génétique standard. En effet, nous préférons éviter que des chromosomes "parents" constituant de bonnes solutions soient remplacés par leurs enfants, ce qui pourrait faire passer l'algorithme à côté de la bonne solution.

Étape 7 : L'objectif de cette étape est de construire la nouvelle génération. En effet, dans notre approche nous employons deux méthodes de sélection. La première est la sélection déterministe qui, dans la population courante composée des parents et des enfants, permet grâce à la notion d'élitisme de trouver un pourcentage $Selection_{det}$ de la nouvelle génération composée des meilleurs chromosomes. Elle est suivie d'une sélection stochastique qui, parmi les chromosomes "parents" et "enfants" restants, complétera la population de façon à ce qu'elle conserve sa taille initiale au cours des générations.

Étape 8 : La dernière étape est un test d'arrêt de l'algorithme. Si un nombre total de générations N_{gen} fixé a priori est atteint, alors on arrête l'évolution de l'AG. Sinon, on répète les étapes 3 à 8. D'autres critères d'arrêt peuvent être utilisés.

³Si $E_{gd} = 1$, alors celle-ci est remplacée par 0 et vice-versa

3.8 Résultat de la mise en correspondance

À présent nous avons mis au point l'ensemble des mécanismes nécessaires au déroulement de l'algorithme génétique. Avant de pouvoir appliquer cet algorithme pour mettre en correspondance une séquence d'images stéréoscopiques réelles, nous devons ajuster l'ensemble des paramètres intervenant dans la fonction d'évaluation ainsi que ceux qui régissent l'évolution de l'algorithme.

3.8.1 Choix des paramètres

Le choix des paramètres est un point important de notre méthode. Comme nous pouvons le constater, celle-ci fait intervenir 2 groupes de paramètres. Le premier groupe est constitué des coefficients de pondération des termes de la fonction d'évaluation (cf. équation 3.5) :

- K_{uni} , le coefficient de pondération du terme correspondant à la contrainte d'unicité ;
- K_{ord} , le coefficient de pondération du terme correspondant à la contrainte d'ordre ;
- K_{cont} , le coefficient de pondération du terme correspondant à la contrainte de continuité de la disparité.
- K_m , le coefficient de pondération du terme favorisant les chromosomes codant un nombre maximal d'appariements.

Le deuxième groupe est composé de cinq paramètres qui sont associés à l'algorithme génétique. Ces paramètres sont :

- N_{gen} , le nombre de générations ;
- $Taille_{pop}$, la taille de la population ;
- P_{crois} , la probabilité de croiser deux individus "parents" tirés au sort par la méthode de la roulette ;
- P_{mut} , la probabilité de muter un gène d'un individu "enfant" issu de l'étape de croisement ;
- $Selection_{det}$ qui représente le pourcentage d'individus de la population courante qui sont sélectionnés de manière déterministe.

3.8.1.1 Choix des coefficients intervenant dans la fonction d'évaluation

La stratégie que nous avons adoptée pour choisir les valeurs des coefficients de pondération des différents termes de la fonction d'évaluation est basée sur l'importance que nous avons

donnée à chacune des contraintes utilisées. Il est évident que les contributions de toutes les contraintes au problème de l'appariement n'ont pas obligatoirement la même importance.

Nous avons choisi de forcer la contribution des contraintes d'unicité, d'ordre et la contribution du terme favorisant un nombre maximal d'appariements par rapport à la contrainte de continuité de la disparité car les objets présents dans les scènes observées peuvent présenter des formes complexes. En tenant comptes de ces remarques, les résultats de notre procédure de mise en correspondance, que nous présentons dans le paragraphe 3.8.2, sont obtenus en utilisant pour les différents coefficients les valeurs suivantes :

- $K_{uni} = K_{ord} = K_m = 5$;
- $K_{cont} = 1$

Il est important de noter que le choix de ces coefficients mérite attention. Il est essentiel aussi de regarder s'il est préférable de définir leurs valeurs en fonction des images à traiter ou bien de les fixer à des valeurs constantes quelque soit le couple d'images stéréoscopiques. Seuls les tests sur des images réelles montreront que l'ajustement proposé permet de traiter des séquences différentes avec succès.

3.8.1.2 Choix des paramètres génétiques

Nous avons vu que cinq paramètres interviennent dans le fonctionnement de notre approche génétique pour l'appariement d'images stéréoscopiques. Nous rappelons également que ces paramètres influent sur l'évolution de la population au cours des générations. En effet, un mauvais choix de ces paramètres pourrait amener la population d'individus à converger vers un minimum local. Après de nombreux tests effectués sur les images-ligne de la séquence "Piéton", nous avons retenu les valeurs suivantes pour les paramètres génétiques :

- $P_{crois} = 0.6$

La probabilité de croisement permet de choisir entre une conservation ou une modification des chromosomes. Pour favoriser l'exploration de l'espace des solutions, la probabilité de croisement est souvent choisie entre 0.5 et 0.9 [Mic94].

- $P_{mut} = 1/\text{taille du chromosome}$

Comme nous l'avons vu précédemment (cf. § 3.5), la taille du chromosome est variable car elle dépend du nombre de couples de points de contour candidats à l'appariement. De ce fait, nous avons opté pour une probabilité de mutation qui s'adapte à la taille du chromosome, à savoir : $P_{mut} = 1/\text{taille du chromosome}$. En effet, une probabilité de

mutation fixe (0.01 par exemple) sur toute la séquence stéréoscopique traitée, peut convenir pour un chromosome de taille 100. Cependant, elle risque d'être inadaptée pour un chromosome de taille plus grande ou plus petite car elle augmente la possibilité d'avoir beaucoup de gènes mutés dans le premier cas ou la diminue dans le deuxième, ce qui pourrait empêcher la convergence de la population vers une solution optimale.

- $Selection_{det} = 0.10$

Nous rappelons que ce paramètre intervient lors de la construction de la nouvelle génération. Grâce à la notion d'élitisme, il permet de sélectionner parmi les chromosomes parents et enfants, les meilleures solutions pour faire partie de la génération suivante. En d'autres termes, pour une population de taille 100, les 10 meilleurs chromosomes de la génération courante sont sélectionnés pour passer à la génération qui suit.

Nous discuterons le choix de valeurs des paramètres N_{gen} et $Taille_{pop}$ lors du traitement de séquences stéréoscopiques.

Dans les paragraphes suivants, nous allons présenter les résultats de mise en correspondance des primitives extraites d'images-ligne de deux séquences stéréoscopiques après les avoir décrites et montré les résultats de l'extraction de leurs contours. Les deux séquences représentent des images stéréoscopiques réelles acquises par le stéréoscope équipé de deux caméras linéaires développé au laboratoire I^3D . La première est la séquence "Piéton" et la deuxième est appelée la séquence "Approche d'un véhicule".

3.8.2 Séquence "Piéton"

3.8.2.1 Description et extraction des points de contour

Dans cette première séquence, la scène représente le déplacement d'un piéton devant le stéréoscope linéaire monté sur un support fixe et visant la chaussée à une distance de 50 mètres (cf. figure 3.17). On rappelle que les axes optiques des deux caméras sont parallèles et écarté d'une distance E et que le plan de visée commun est légèrement incliné par rapport à l'horizontale d'un angle θ . Du fait de cette inclinaison, tout obstacle, même de faible hauteur, qui entre dans ce secteur de stéréovision est vu simultanément par les caméras gauche et droite. Il peut ainsi être détecté et localisé par triangulation.

Les caractéristiques du stéréoscope linéaire sont les suivantes :

- Définition des caméras : 1728 pixels.

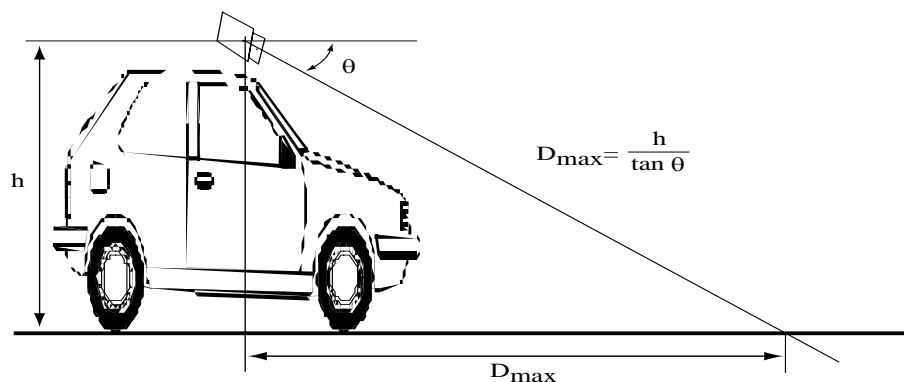


FIG. 3.17 : Configuration du stéréoscope linéaire (Vue de côté).

- Largeur des capteurs : 22,1 mm.
- Distance focale des objectifs : 100 mm.
- Entre axe : $E = 1\text{m}$.

Pour améliorer la lisibilité, des images-ligne sont enregistrées sous forme de séquences. Même si elle est parfois difficile à interpréter, la visualisation d'une séquence permet, avec une certaine expérience, d'avoir une idée de la scène qui a été observée. De plus, une séquence stockée sous la forme d'un fichier est plus facile à manipuler que plusieurs fichiers contenant chacun une seule image-ligne de la séquence enregistrée.

Pour réaliser les enregistrements, un disque dur d'un PC est utilisé comme un magnéto-scope, et le stéréoscope comme une caméra vidéo. La visualisation est, quant à elle, réalisée grâce au moniteur VGA du PC hôte pour vérifier la qualité des enregistrements. L'acquisition des images gauche et droite est réalisée en interfaçant les deux caméras avec une seule carte d'acquisition linéaire, par l'intermédiaire d'un multiplexeur développé spécifiquement pour garantir la simultanéité des prises de vue stéréoscopiques. Le rôle du multiplexeur est de gérer les signaux de synchronisation et de commande des deux caméras et de la carte d'acquisition, ainsi que de contrôler le transfert des pixels vers cette dernière.

La figure 3.18 décrit schématiquement les déplacements du piéton au cours du temps. On le voit d'abord se diriger vers son point de départ (B), situé un peu au-delà de l'intersection du plan de visée avec la chaussée (A). Ce faisant, il passe hors de portée du stéréoscope linéaire et, en conséquence, disparaît de l'image. 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 vision stéréoscopique (E).

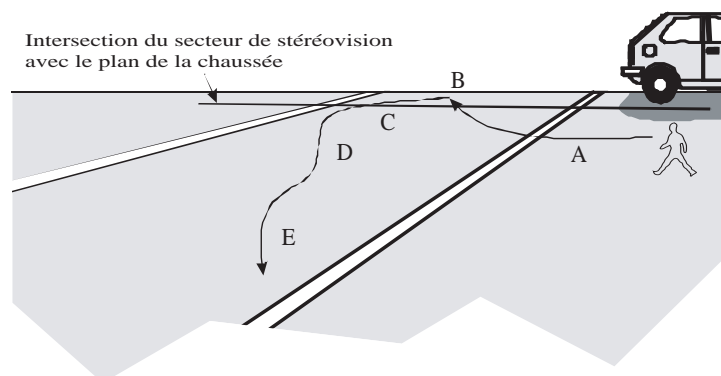


FIG. 3.18 : Description de la scène "Piéton".

Les séquences gauche et droite issues des deux caméras sont représentées sous la forme d'images bidimensionnelles en "empilant" les différentes images-ligne acquises (cf. figure 3.19).

Les séquences se lisent comme suit :

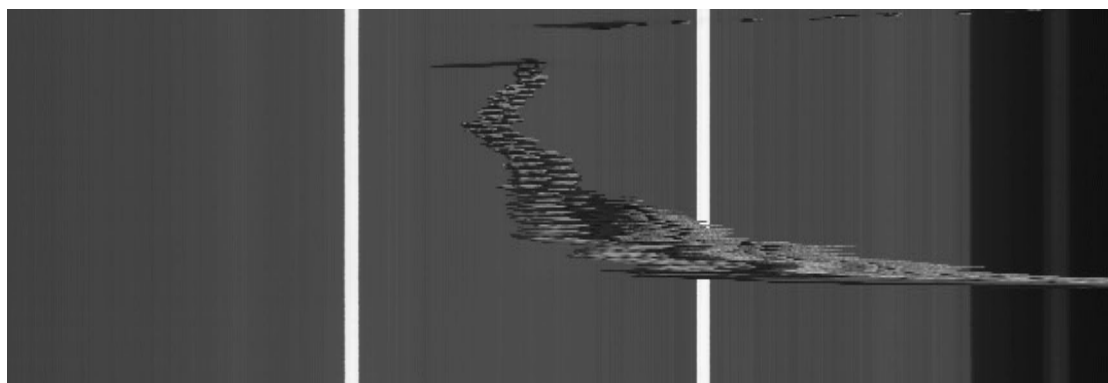
- L'axe horizontal représente le plan de visée du stéréoscope. Chaque ligne correspond donc à une image-ligne.
- L'axe vertical représente l'axe du temps, dirigé vers le bas de l'image. Chacune des deux séquences est composée de 200 images-ligne. L'intervalle de temps entre deux prises de vue consécutives est de 100 ms.

En plus du déplacement du piéton, sur les images de la séquence, on voit clairement les lignes blanches de la chaussée ainsi que l'ombre projetée par une voiture située hors du champ du stéréoscope. Cette ombre est visible sur la droite des images comme une zone entièrement noire (cf. figure 3.19).

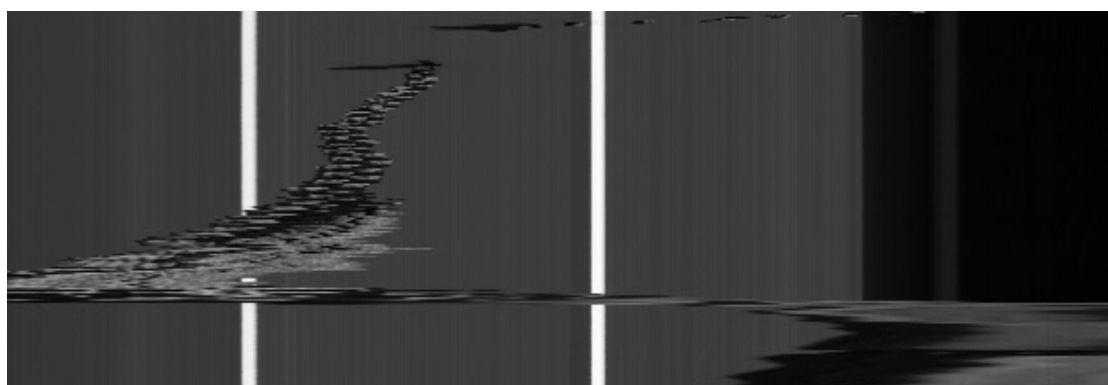
La figure 3.20 indique les points de contour résultants de l'application du module d'extraction des primitives décrit au § 2.6.2.3. Pour obtenir une meilleure lisibilité, les points de contour sont représentés par des points noirs, et ceux quelque soit la valeur de l'amplitude de leur dérivée.

On constate une bonne détection des points de contour. Les bords des lignes blanches ainsi que les frontières entre la route et la zone d'ombre sont parfaitement détectés. De plus, aucune fausse détection n'apparaît au niveau de la chaussée malgré la présence de certaines irrégularités.

Il faut noter que le piéton génère de nombreux points de contour. L'ensemble de ces points forme une structure cohérente au cours du temps. Mais les points de contour pris séparément ne respectent pas cette cohérence temporelle, car un point de contour d'une image de la séquence



(a) Séquence gauche.



(b) Séquence droite.

FIG. 3.19 : Déplacement d'un piéton vu par le stéréoscope linéaire.

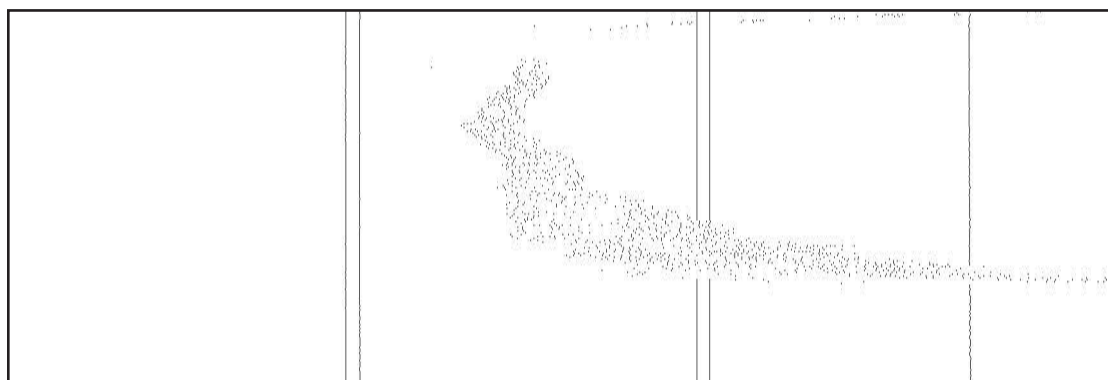
n'a pas obligatoirement de correspondant dans l'image suivante.

3.8.2.2 Résultat de la mise en correspondance

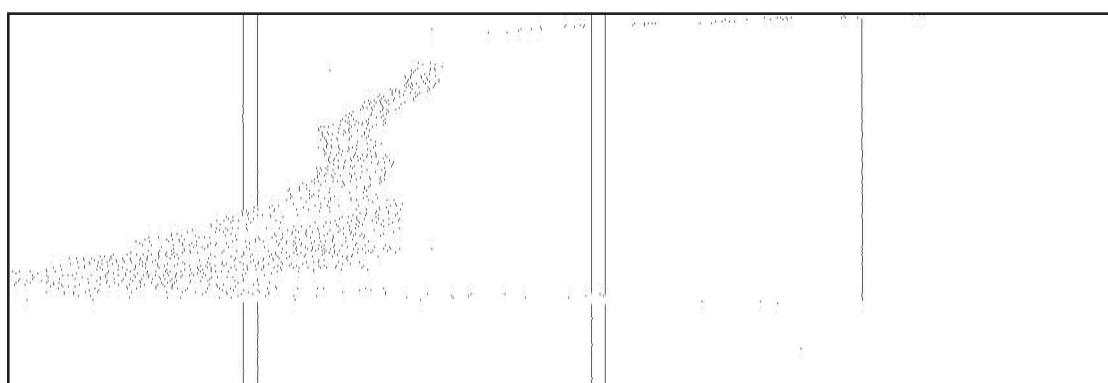
La procédure génétique de mise en correspondance décrite au paragraphe 3.7 est appliquée à tous les couples d'images-ligne épipolaires de la séquence "Piéton". Chaque couple d'images-ligne est traité indépendamment des autres. À l'issue de chaque traitement, une liste des paires de points de contour mis en correspondance est constituée. Chaque paire de points de contour appariés permet, grâce aux formules de triangulation, de déterminer la distance et la position d'un point de la scène réelle correspondant.

Le résultat de mise en correspondance présenté sur la figure 3.21 est obtenu avec $Taille_{pop} = 50$ et $N_{gen} = 300$. Ces valeurs ont été choisies expérimentalement après avoir effectué plusieurs essais sur cette séquence stéréoscopique. Le résultat obtenu s'interprète de la manière suivante :

- L'axe horizontal représente l'axe des abscisses de la scène. Un point se trouvant sur le



(a) Points de contour de la séquence gauche.



(b) Points de contour de la séquence droite.

FIG. 3.20 : Extraction des contours de la séquence "Piéton".

bord gauche (resp. droit) de l'image est situé à $5m$ à gauche (resp. à droite) de l'axe optique.

- L'axe vertical représente l'axe du temps, celui-ci défilant du haut vers le bas de l'image.
- La distance est représentée en couleur. Celle-ci varie du bleu, qui correspond à la distance la plus grande par rapport au stéréoscope, au rouge qui correspond à la distance la plus petite.

La figure 3.21 montre que les points de contour des bandes ainsi que la transition entre la chaussée et la zone d'ombre à droite de la figure sont bien appariés et leur détection est stable tout au long de la séquence. En effet, les positions et les distances calculées sont constantes de ligne en ligne. La couleur bleue indique que les bandes ainsi que la transition entre la chaussée et la zone d'ombre sont les plus éloignées du stéréoscope. Le piéton est également bien détecté. En effet, la couleur des points de contour correspondant à celui-ci varie du bleu au rouge, ce qui indique son rapprochement du stéréoscope.

On remarque cependant sur le résultat obtenu la présence de mauvais appariements lorsque

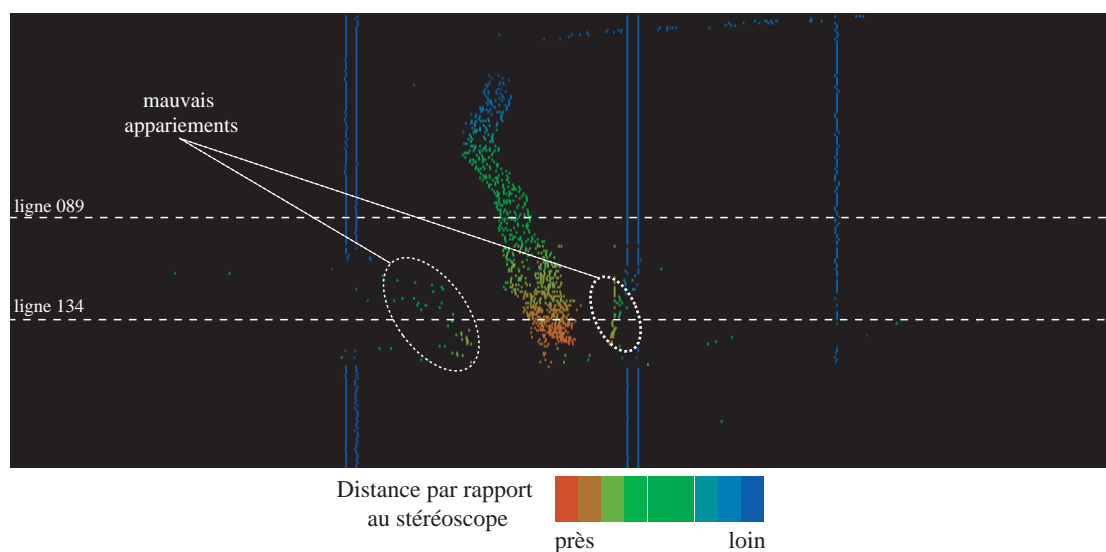


FIG. 3.21 : Résultat de la mise en correspondance de la séquence stéréoscopique "Piéton" grâce au codage binaire avec 50 chromosomes et 300 générations.

des occlusions se produisent, c'est-à-dire lorsque le piéton cache l'une des bandes blanches à l'une des caméras. Ces appariements erronés sont provoqués par la mise en correspondance des points de contour d'une des bandes blanches avec ceux du piéton (*cf.* figure 3.21).

Afin d'évaluer le comportement de l'algorithme génétique de mise en correspondance lors de la procédure d'appariement des points de contour détectés sur les images-ligne stéréoscopiques de la séquence "Piéton", nous montrons les courbes de l'évolution de la valeur de la fonction d'évaluation pour le meilleur chromosome et ceci pour deux couples d'images-ligne épipolaires. Le premier couple est celui de rang 89. Ce couple est sélectionné car d'une part, le nombre de points de contour détectés dans chacune des images-ligne épipolaires gauche et droite n'est pas très élevé, à savoir 13 points (*cf.* figure 3.22), et d'autre part, ce couple est situé dans une zone où il n'y a pas d'occlusions (le piéton, les bandes blanches et la transition entre la chaussée et la zone d'ombre sont visibles par les deux caméras). Le deuxième couple sélectionné est celui de rang 134. Contrairement au premier couple, celui-ci contient un nombre important de points de contour (29 dans chacune des images-ligne épipolaires (*cf.* figure 3.22)) et il coïncide avec le moment où le piéton cache l'une des bandes blanches à l'une des caméras.

Les courbes de la figure 3.23 montrent que pour le couple d'images-ligne de rang 89, la valeur de la fonction d'évaluation atteinte au bout de 210 générations est voisine de 2 et qu'aucune amélioration est constatée au cours des 90 générations suivantes.

En ce qui concerne le couple d'images-ligne de rang 134 mettant en œuvre un nombre re-

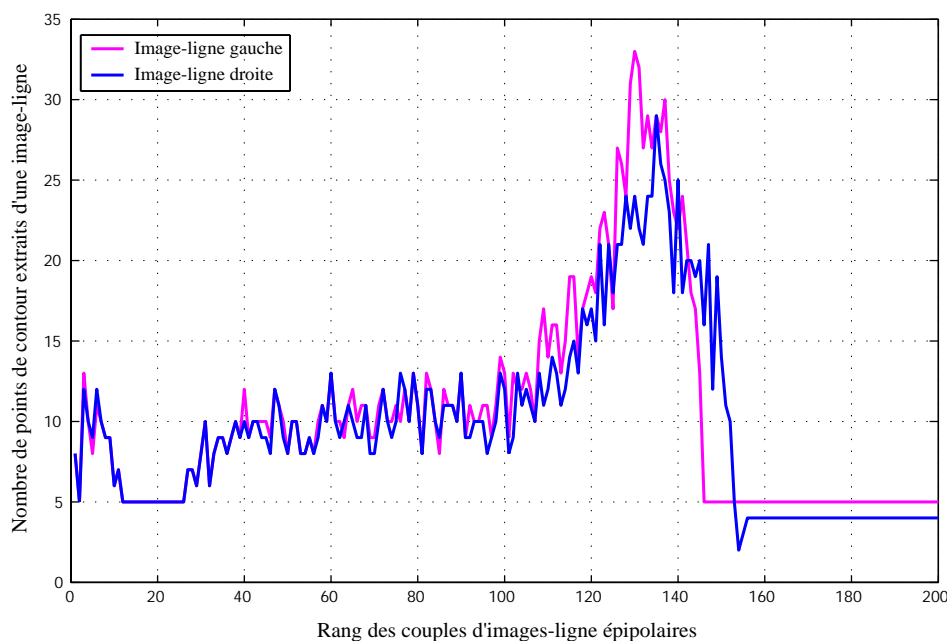


FIG. 3.22 : Nombre de points de contour extraits des images-ligne de la séquence "Piéton".

lativement important de points de contour, la figure 3.24 montre que la valeur de la fonction d'évaluation du meilleur chromosome évolue d'une manière permanente et que celle-ci tombe vers 835 au bout de 300 générations. En d'autres termes, ces figures nous permettent de conclure que les valeurs des paramètres $Taille_{pop}$ et N_{gen} sont convenables pour l'appariement des primitives de couples d'images-ligne contenant un nombre limité de points de contour et que des valeurs plus élevées se révèlent nécessaires pour des couples d'images-ligne plus complexes en terme du nombre de points de contour à apparier.

Pour améliorer les mises en correspondance des primitives des images-ligne épipolaires de la séquence "Piéton", nous traitons de nouveau cette séquence en augmentant les valeurs de deux paramètres génétiques, à savoir la taille $Taille_{pop}$ de la population et le nombre N_{gen} de générations. La figure 3.25 montre les résultats de la mise en correspondance obtenus pour différentes valeurs de $Taille_{pop}$ et N_{gen} .

Les courbes des figures 3.26, 3.27, 3.28, 3.29, 3.30 et 3.31 montrent l'évolution de la valeur de la fonction d'évaluation pour le meilleur chromosome lors du déroulement de l'algorithme génétique de mise en correspondance des primitives des images-ligne épipolaires gauche et droite du couple de rang 89 ainsi que celui de rang 134 pour différentes valeurs de $Taille_{pop}$ et de N_{gen} . Ces courbes montrent que pour le couple de rang 89, la meilleure valeur atteinte par la fonction d'évaluation reste la même, à savoir 2, quelque soient les valeurs des para-

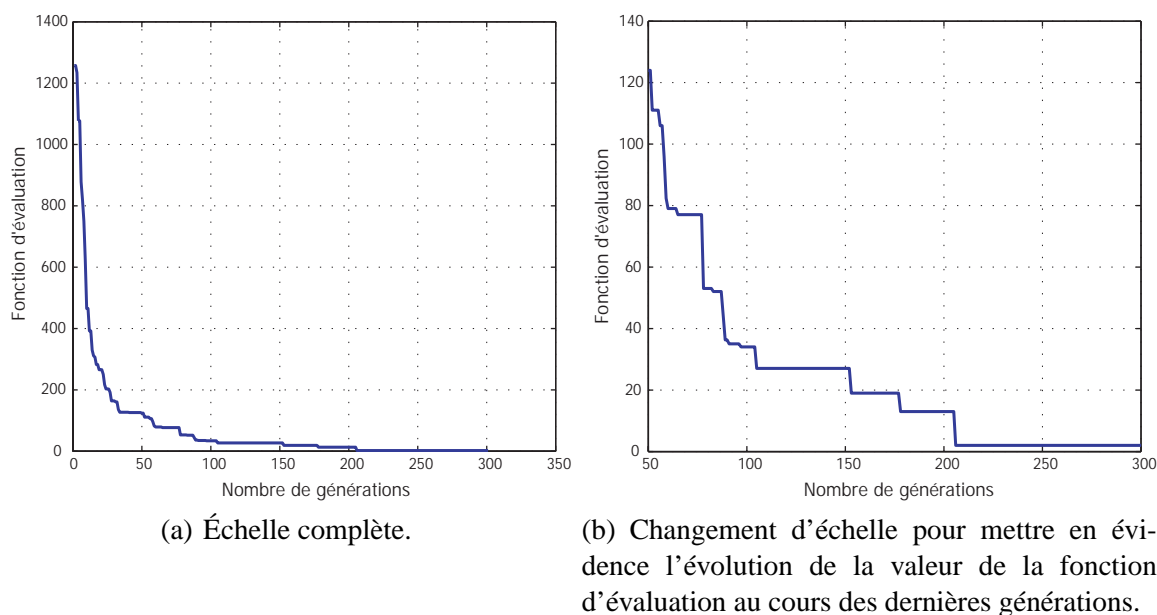


FIG. 3.23 : Évolution de la fonction d'évaluation pour le meilleur chromosome avec $Taille_{pop} = 50$ et $N_{gen} = 300$ pour le couple d'images-ligne épipolaires de rang 89 de la séquence "Piéton".

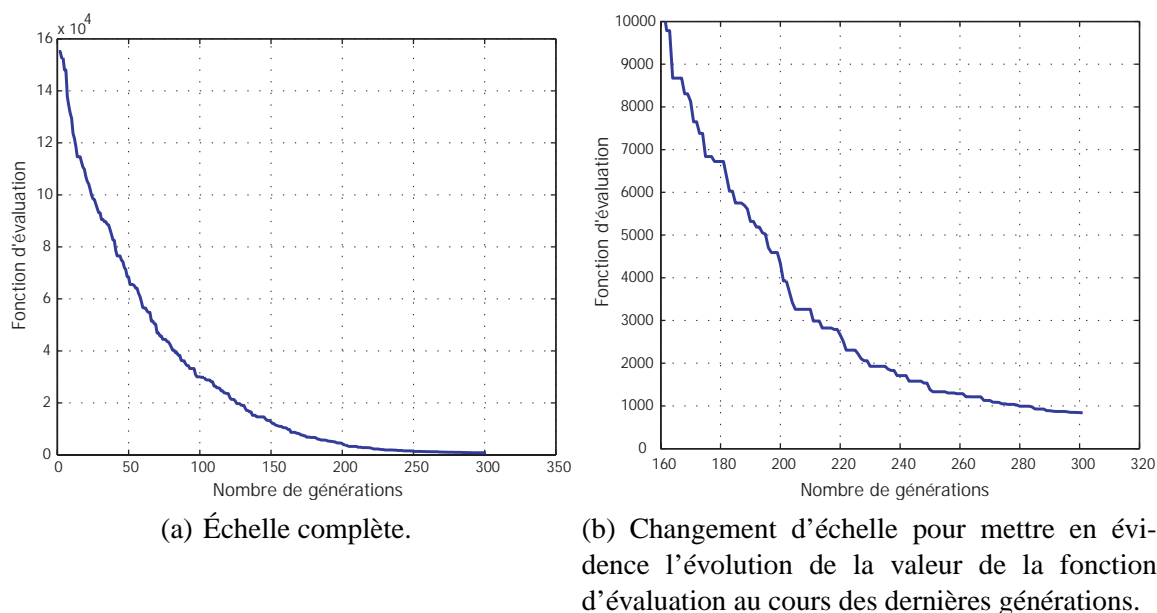
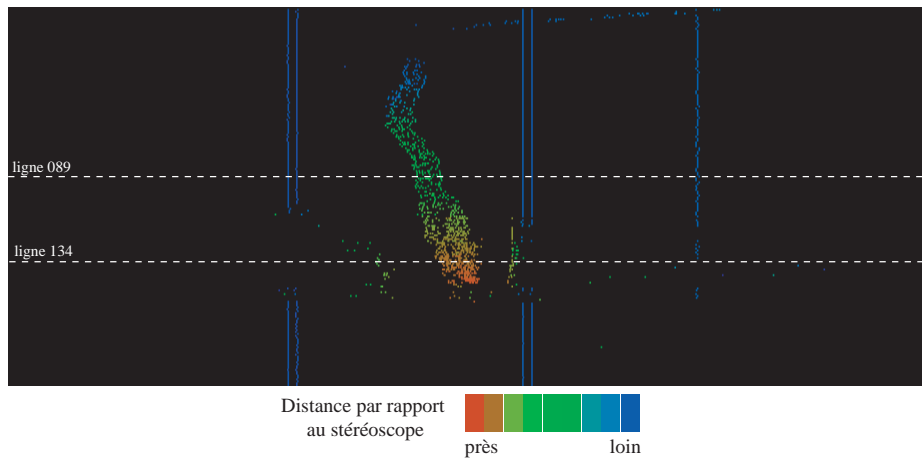
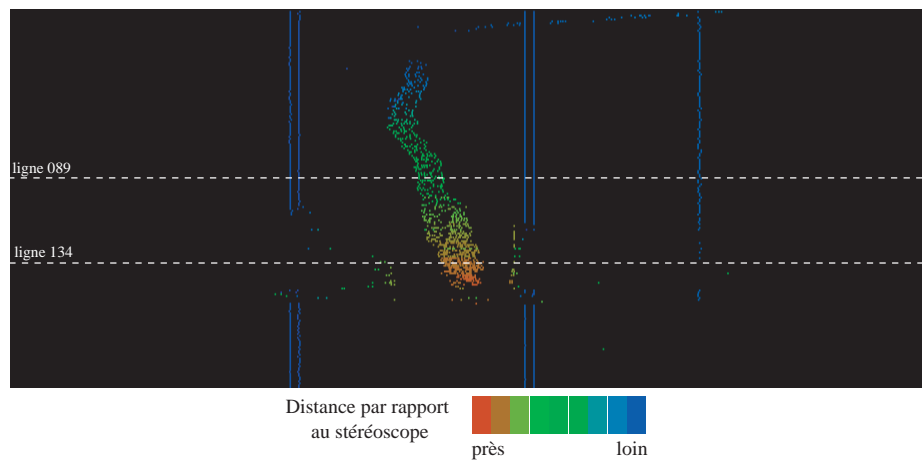
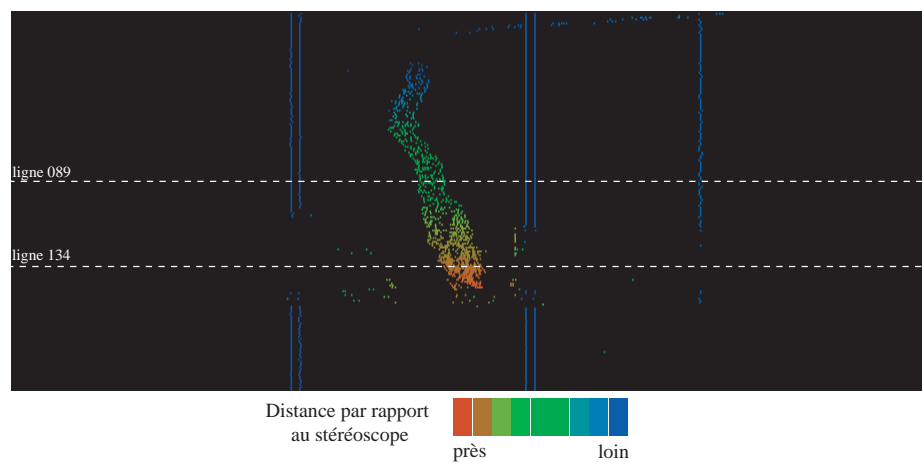


FIG. 3.24 : Évolution de la fonction d'évaluation pour le meilleur chromosome avec $Taille_{pop} = 50$ et $N_{gen} = 300$ pour le couple d'images-ligne épipolaires de rang 134 de la séquence "Piéton".

mètres $Taille_{pop}$ et N_{gen} (cf. figures 3.26(b), 3.27(b) et 3.28(b)). Pour le couple de rang 134, la meilleure valeur atteinte est de -267 et ceci lorsque $Taille_{pop} = 100$ et $N_{gen} = 1000$ (cf. figures 3.29(b), 3.30(b) et 3.31(b)).

En comparant les courbes des figures 3.26(b), 3.27(b) et 3.28(b) avec la courbe de la fi-

(a) $Taille_{pop} = 100$ et $N_{gen} = 300$.(b) $Taille_{pop} = 50$ et $N_{gen} = 1000$.(c) $Taille_{pop} = 100$ et $N_{gen} = 1000$.**FIG. 3.25 :** Résultat de la mise en correspondance de la séquence stéréoscopique "Piéton" pour différentes valeurs de $Taille_{pop}$ et de N_{gen} .

gure 3.26(a), nous constatons que pour le couple de rang 89, l'augmentation des valeurs des paramètres $Taille_{pop}$ et N_{gen} n'a pas permis d'atteindre une solution meilleure que celle déjà obtenue précédemment. Mais en comparant les courbes des figures 3.29(b), 3.30(b) et 3.31(b) avec la courbe de la figure 3.24(b), nous constatons que pour le couple de rang 134 mettant en œuvre un nombre important de points de contour candidats à l'appariement, l'augmentation des valeurs des paramètres $Taille_{pop}$ et N_{gen} a permis à l'algorithme génétique de mieux explorer l'espace de recherche et ainsi de converger vers une solution meilleure. En effet, la meilleure valeur de la fonction d'évaluation atteinte lors d'appariement du couple de rang 134 est de -267 pour $Taille_{pop} = 100$ et $N_{gen} = 1000$ à la place de 835 pour $Taille_{pop} = 50$ et $N_{gen} = 300$.

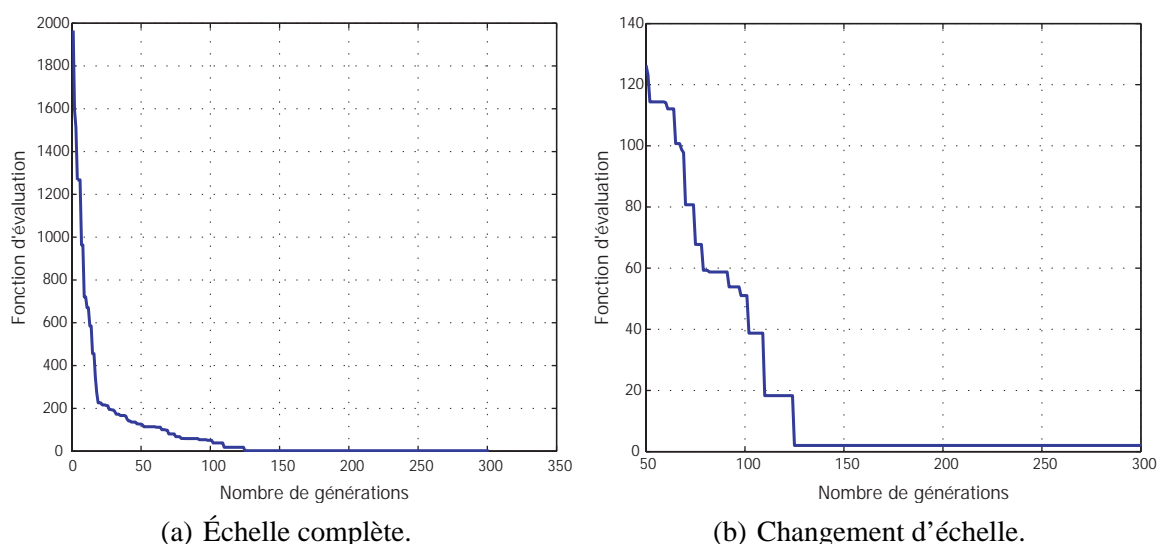


FIG. 3.26 : Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 100$ et $N_{gen} = 300$ pour le couple de rang 89 de la séquence "Piéton". (Le changement d'échelle est effectué pour mettre en évidence l'évolution de la fonction d'évaluation au cours des dernières générations).

Afin de comparer plus concrètement les résultats de mise en correspondance des primitives des couples d'images-ligne de la séquence "Piéton" et ceci pour les différentes valeurs de $Taille_{pop}$ et de N_{gen} , les meilleures valeurs de la fonction d'évaluation atteintes pour chacun de ces couples sont indiquées sur les figures 3.32 et 3.33.

Ces courbes montrent que lorsqu'il s'agit d'apparier les points de contour de couples d'images-ligne contenant un nombre limité de ceux-ci (entre 0 et 15 points de contour environ pour chaque image-ligne), les valeurs de la fonction d'évaluation atteintes sont pratiquement les mêmes et ceci pour les différentes valeurs utilisées pour $Taille_{pop}$ et N_{gen} (voir les valeurs correspondant aux couples compris entre les rangs 1 et 110 ainsi qu'aux couples compris entre les rangs

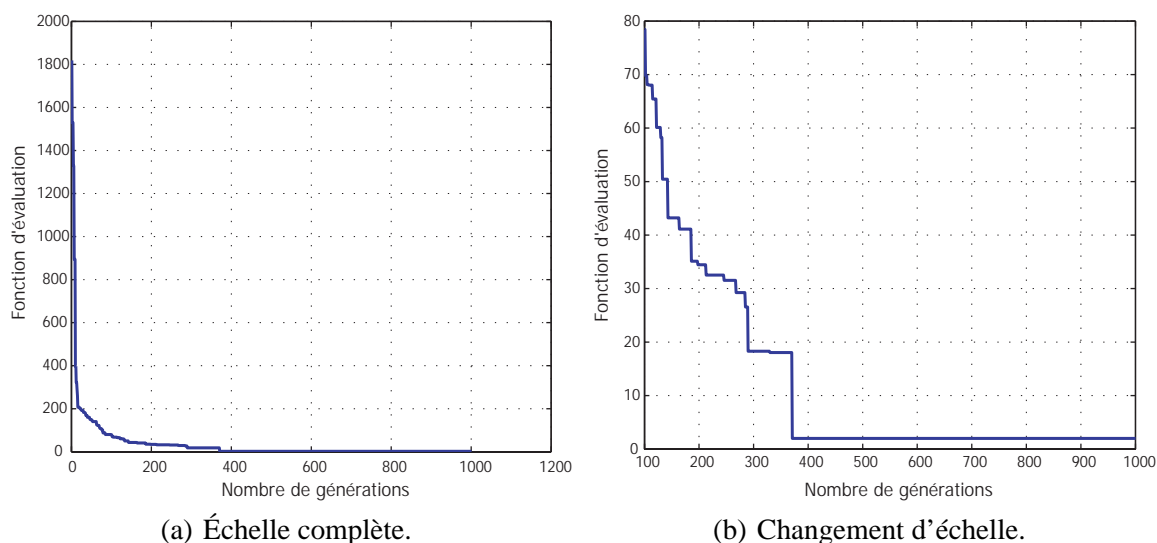


FIG. 3.27 : Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 50$ et $N_{gen} = 1000$ pour le couple de rang 89 de la séquence "Piéton". (Le changement d'échelle est effectué pour mettre en évidence l'évolution de la fonction d'évaluation au cours des dernières générations).

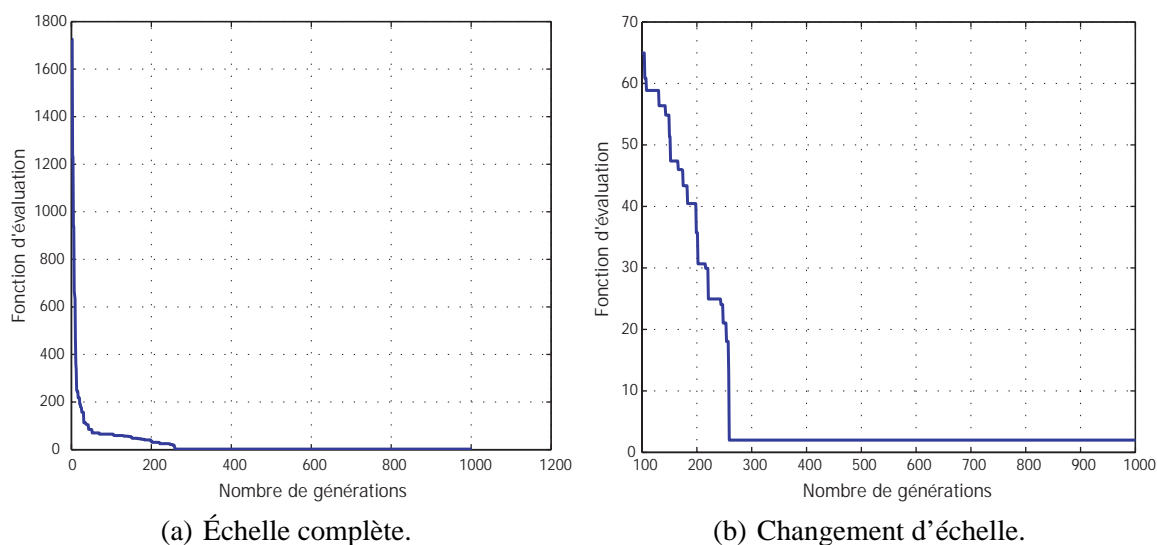


FIG. 3.28 : Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 100$ et $N_{gen} = 1000$ pour le couple de rang 89 de la séquence "Piéton". (Le changement d'échelle est effectué pour mettre en évidence l'évolution de la fonction d'évaluation au cours des dernières générations).

150 et 200). Mais, quand ce un nombre devient important (plus que 15), les courbes montrent que globalement, les meilleures valeurs de la fonction d'évaluation sont celles obtenues pour $Taille_{pop} = 100$ et $N_{gen} = 1000$. En effet, si nous intéressons tout particulièrement à deux valeurs qui sont les valeurs maximale et minimale de chacune de ces courbes, nous constatons que celles-ci sont égales approximativement à :

$$850 \text{ et } -100 \text{ pour } Taille_{pop} = 50 \text{ et } N_{gen} = 300,$$

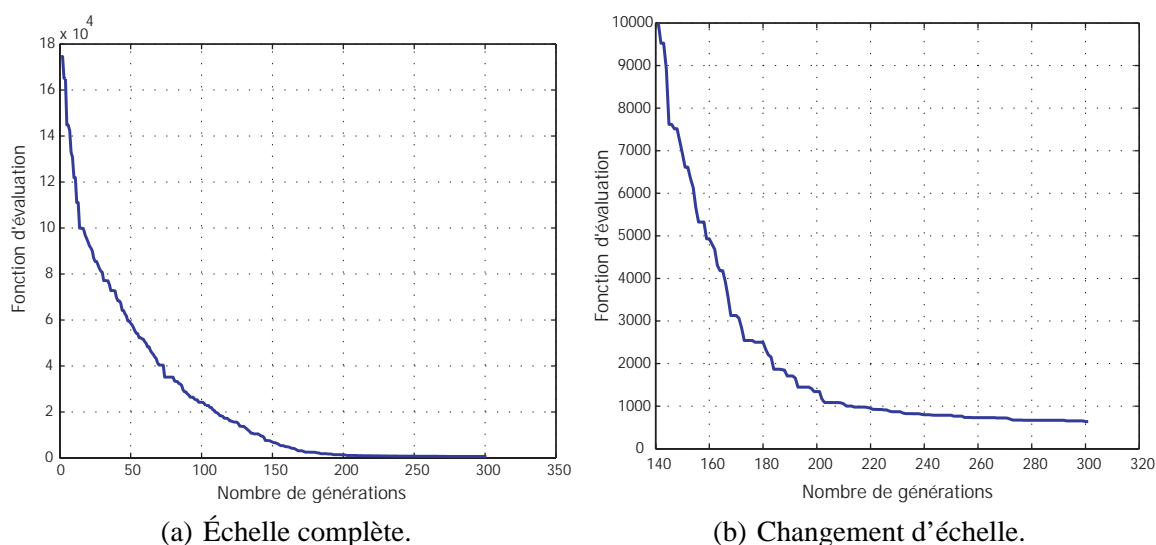


FIG. 3.29 : Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 100$ et $N_{gen} = 300$ pour le couple de rang 134 de la séquence "Piéton". (Le changement d'échelle est effectué pour mettre en évidence l'évolution de la fonction d'évaluation au cours des dernières générations).

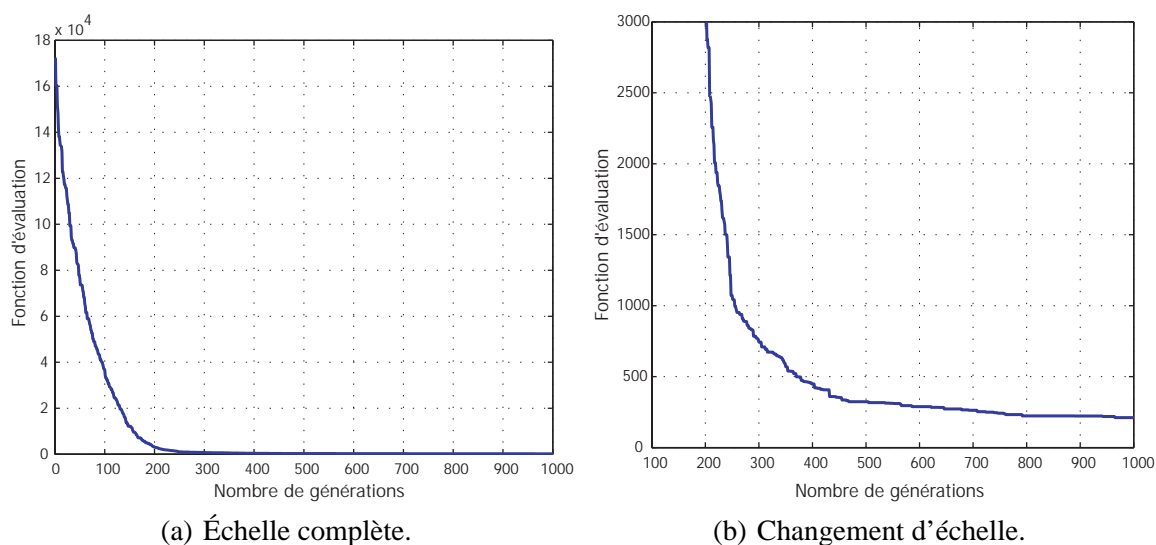


FIG. 3.30 : Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 50$ et $N_{gen} = 1000$ pour le couple de rang 134 de la séquence "Piéton". (Le changement d'échelle est effectué pour mettre en évidence l'évolution de la fonction d'évaluation au cours des dernières générations).

215 et -100 pour $Taille_{pop} = 50$ et $N_{gen} = 1000$,

600 et -100 pour $Taille_{pop} = 100$ et $N_{gen} = 300$,

200 et -280 pour $Taille_{pop} = 100$ et $N_{gen} = 1000$.

En d'autres termes, la courbe ayant les valeurs maximale et minimale les plus petites est celle qui correspond à $Taille_{pop} = 100$ et $N_{gen} = 1000$ (cf. figures 3.32 et 3.33).

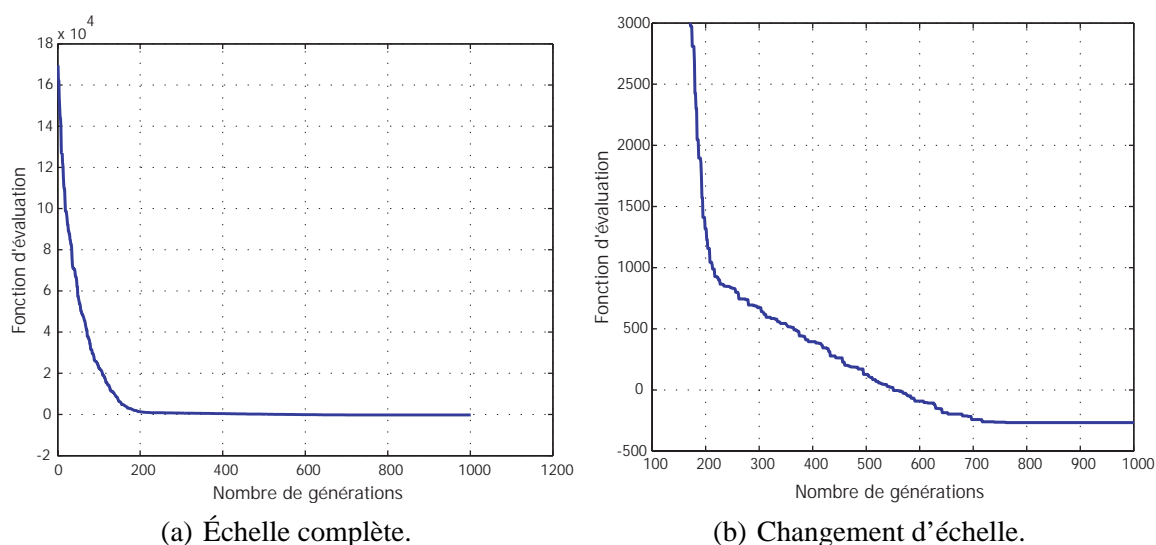


FIG. 3.31 : Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 100$ et $N_{gen} = 1000$ pour le couple de rang 134 de la séquence "Piéton". (Le changement d'échelle est effectué pour mettre en évidence l'évolution de la fonction d'évaluation au cours des dernières générations).

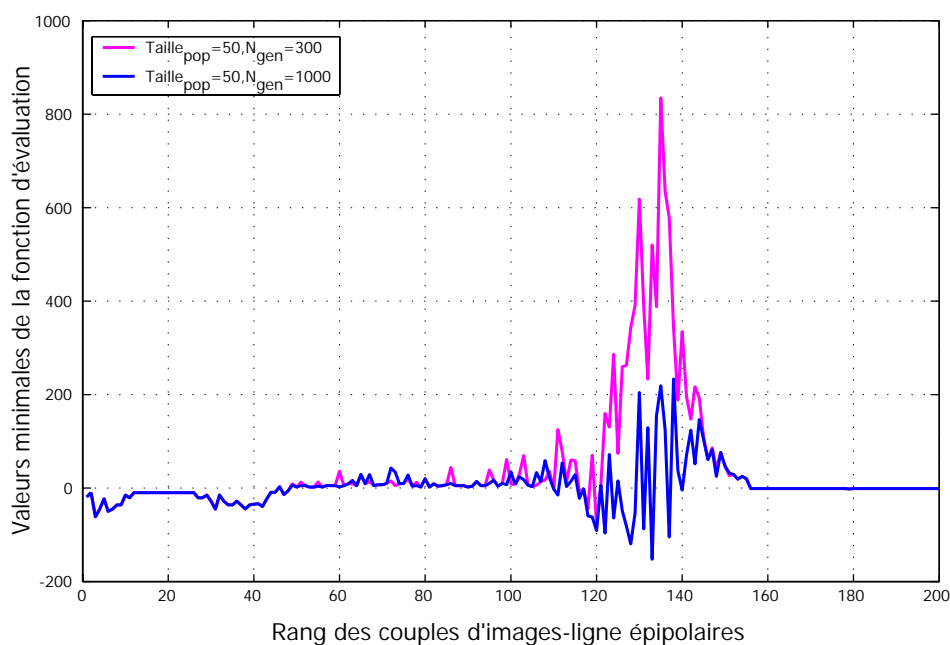


FIG. 3.32 : Valeurs minimales de la fonction d'évaluation pour chacun des couples d'images-ligne de la séquence "Piéton" avec d'une part $Taille_{pop} = 50$ et $N_{gen} = 300$, et d'autre part $Taille_{pop} = 50$ et $N_{gen} = 1000$.

3.8.2.3 Conclusion

Les tests effectués sur la séquence "Piéton" montrent que les résultats de mise en correspondance obtenus sont satisfaisants lorsqu'il s'agit d'apparier des points de contour d'images-ligne stéréoscopiques en contenant un nombre relativement limité. Lorsque le nombre de ces points

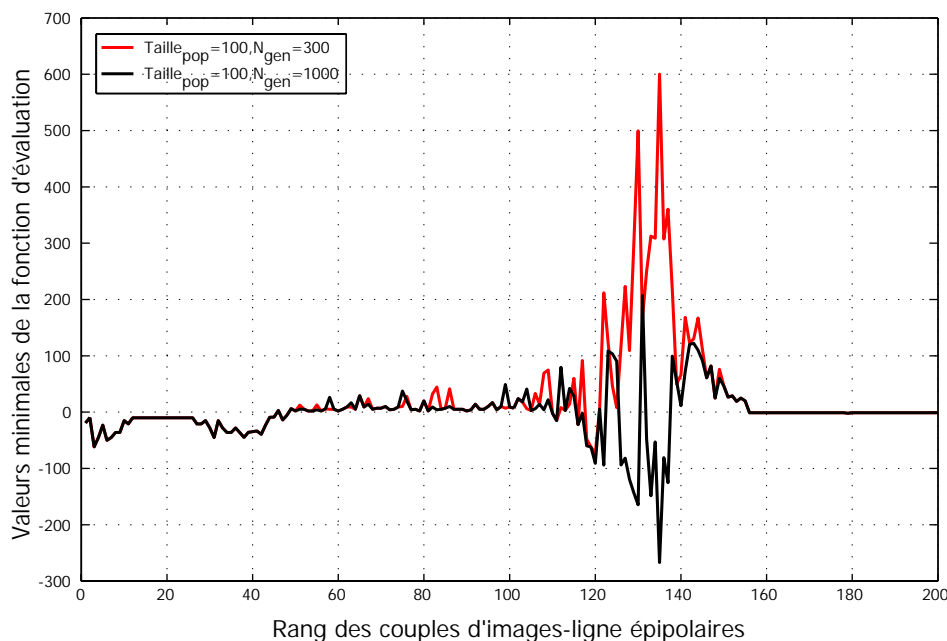


FIG. 3.33 : Valeurs minimales de la fonction d'évaluation pour chacun des couples d'images-ligne de la séquence "Piéton" avec d'une part $Taille_{pop} = 100$ et $N_{gen} = 300$, et d'autre part $Taille_{pop} = 100$ et $N_{gen} = 1000$.

augmente, il est nécessaire d'augmenter le nombre de générations et la taille de la population. Le tableau suivant indique le temps nécessaire pour obtenir le résultat de la mise en correspondance des primitives extraites de cette séquence composée de 200 images-ligne stéréoscopiques et ceci pour différentes valeurs de $Taille_{pop}$ et de N_{gen} (cf. figure 3.34). Ce tableau montre que l'augmentation de la taille de la population et du nombre de générations implique une augmentation considérable du temps de traitement. En effet, pour $Taille_{pop} = 50$ et $N_{gen} = 300$ le temps du traitement est de l'ordre de 17 minutes (c'est-à-dire 5 secondes par couple d'images-ligne) et il passe à environ 74 minutes (c'est-à-dire 22 secondes par couple d'images-ligne) pour $Taille_{pop} = 100$ et $N_{gen} = 1000$. Ces valeurs correspondent à des traitements effectués avec un PC Intel-Pentium IV-2GHz.

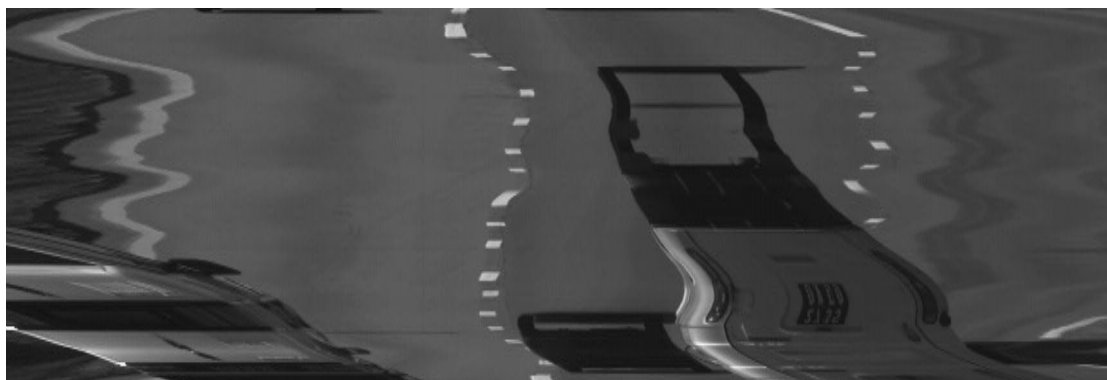
Temps de traitement [sec]		N_{gen}	
		300	1000
$Taille_{pop}$	50	1015	2335
	100	1750	4437

FIG. 3.34 : Temps de traitement de la séquence "Piéton" pour différentes valeurs de $Taille_{pop}$ et de N_{gen} .

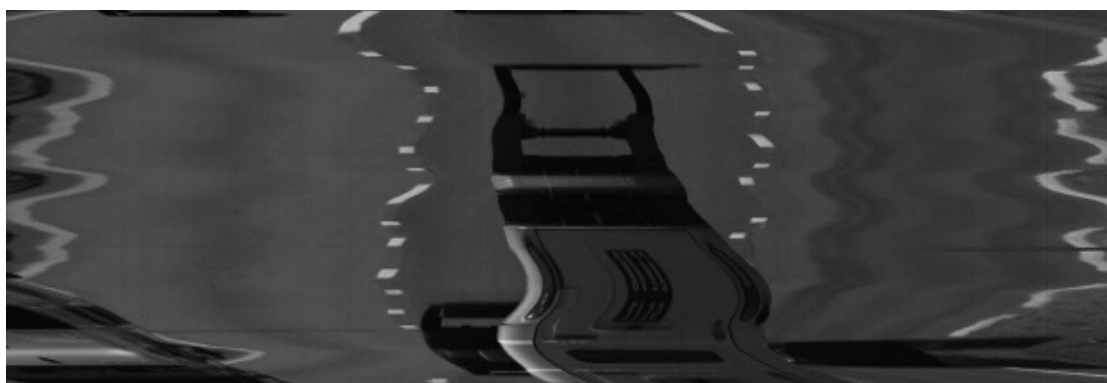
3.8.3 Séquence "Approche d'un véhicule"

3.8.3.1 Description et extraction des points de contour

Dans ce paragraphe nous allons appliquer le module d'extraction des points de contour décrit au § 2.6.2.3, sur tous les couples d'images-ligne épipolaires de la séquence "Approche d'un véhicule", présentée sur la figure 3.35.



(a) Séquence gauche.



(b) Séquence droite.

FIG. 3.35 : Séquence stéréoscopique "Approche d'un véhicule".

Afin d'aider le lecteur à l'interprétation de la séquence, la figure 3.36 donne une description schématique de la scène filmée.

Le véhicule prototype roule derrière le véhicule (A) en se rapprochant de celui-ci. Le plan de visée croise l'arrière du véhicule (A) à une hauteur croissante à mesure que la distance décroît, ce qui explique l'inversion de l'image du véhicule (A) dont on commence à voir les roues pour remonter jusqu'au toit.

Devant le véhicule (A) se trouve un autre véhicule (B) en train de se rabattre après l'avoir dépassé. Par ailleurs, le véhicule prototype est lui-même dépassé par un autre véhicule (C) qui apparaît à la fin de la séquence mais qui n'entre pas entièrement dans le secteur de stéréovision.

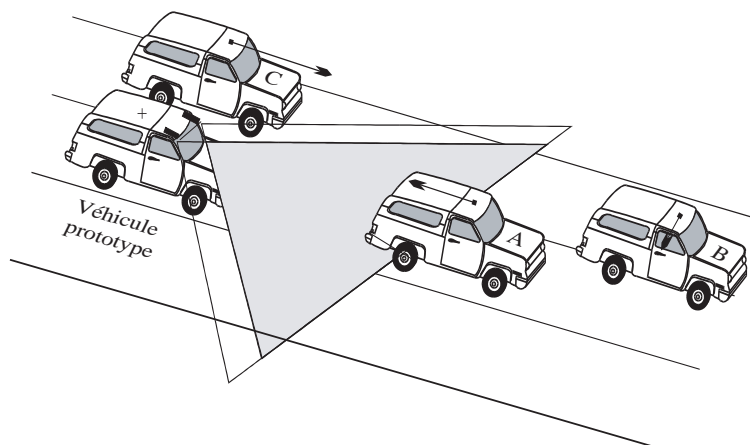


FIG. 3.36 : Déplacements des différents véhicules au cours de la séquence. Les flèches indiquent les mouvements relatifs des véhicules par rapport au véhicule prototype.

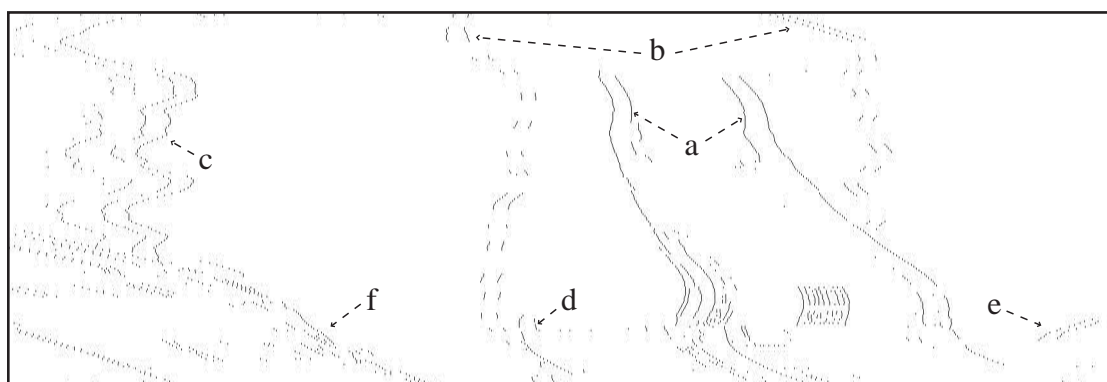
Enfin, on remarque le mouvement de tangage du véhicule prototype, visible surtout au début de la séquence sous la forme d'une variation de la largeur de l'image de la chaussée.

Comme précédemment, chaque couple d'images-ligne épipolaires est traité indépendamment des autres. La figure 3.37 illustre les points de contour extraits des séquences gauche et droite. Comme pour la séquence "Piéton", afin d'obtenir une bonne lisibilité, les points de contour sont représentés par des points noirs quelque soit la valeur de l'amplitude de leur dérivée.

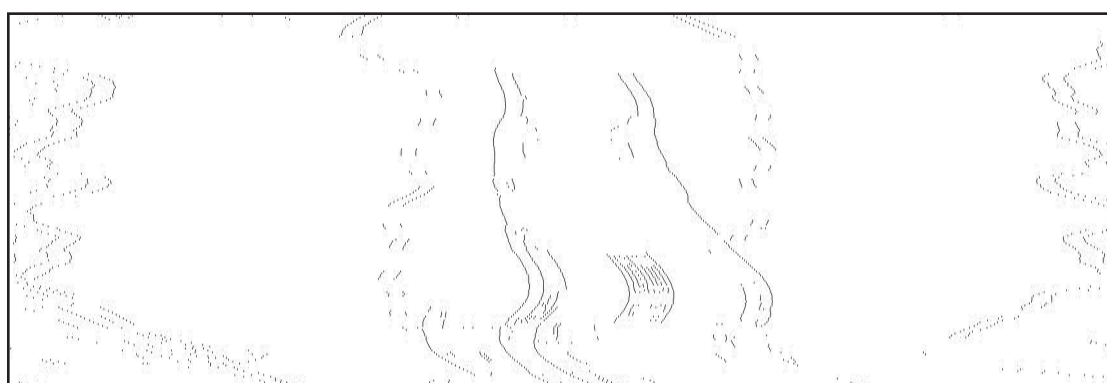
Sur cette figure on aperçoit les points de contour (a) correspondant au rapprochement du véhicule (A) et du véhicule prototype, au cours du temps. On voit également, en pointillés, les deux bandes qui séparent les voies de l'autoroute (b). Leur ondulation est due au mouvement de lacet du véhicule prototype par rapport à la chaussée. La ligne qui délimite le terre-plein central est également détectée, du moins lorsqu'elle est dans le champ des deux caméras (c). Les points de contour qui correspondent au véhicule (B) qui se rabat devant le véhicule (A) sont visibles en bas de l'image (d). Ces points correspondent à son extrémité gauche. Son extrémité droite n'est pas visible en raison d'une occlusion par le véhicule (A) qui la masque à la caméra droite. L'extrémité de sa longue ombre est visible en bas, à l'extrême droite de l'image (e). Le véhicule qui double le véhicule prototype est visible, en bas et à gauche de l'image (f).

3.8.3.2 Résultat de la mise en correspondance

La procédure génétique de mise en correspondance décrite au paragraphe 3.7 est appliquée à tous les couples d'images-ligne épipolaires de la séquence "Approche d'un véhicule" présentés



(a) Points de contour de la séquence gauche.



(b) Points de contour de la séquence droite.

FIG. 3.37 : Extraction des points de contours de la séquence "Approche d'un véhicule".

dans le paragraphe 3.8.3.1.

Comme précédemment, chaque couple d'images-ligne est traité indépendamment des autres. A l'issue de chaque traitement, une liste des paires de points de contour mis en correspondance est constituée. Chaque paire de points de contour appariés permet, grâce aux formules de triangulation, de déterminer la distance et la position d'un point de la scène réelle.

Les valeurs des paramètres génétiques choisies pour traiter cette séquence sont identiques à celles déjà utilisées lors du traitement de la séquence stéréoscopique "Piéton" décrite au paragraphe 3.8.1.2.

Les résultats de mise en correspondance obtenus avec les différentes valeurs des paramètres $Taille_{pop}$ et N_{gen} , sont très proches les uns des autres. Ces résultats sont montrés sous forme de courbes (cf. figures 3.40 et 3.41). L'un d'entre eux, seulement, correspondant au résultat obtenu avec $Taille_{pop} = 100$ et $N_{gen} = 300$, est présenté sous la forme d'une image (cf. figure 3.38). Ce résultat est choisi car des valeurs identiques pour $Taille_{pop}$ et N_{gen} seront utilisées pour mettre en correspondance les primitives de cette séquence, mais avec des AGs utilisant une autre type

de codage génétique (cf. § 4.8).

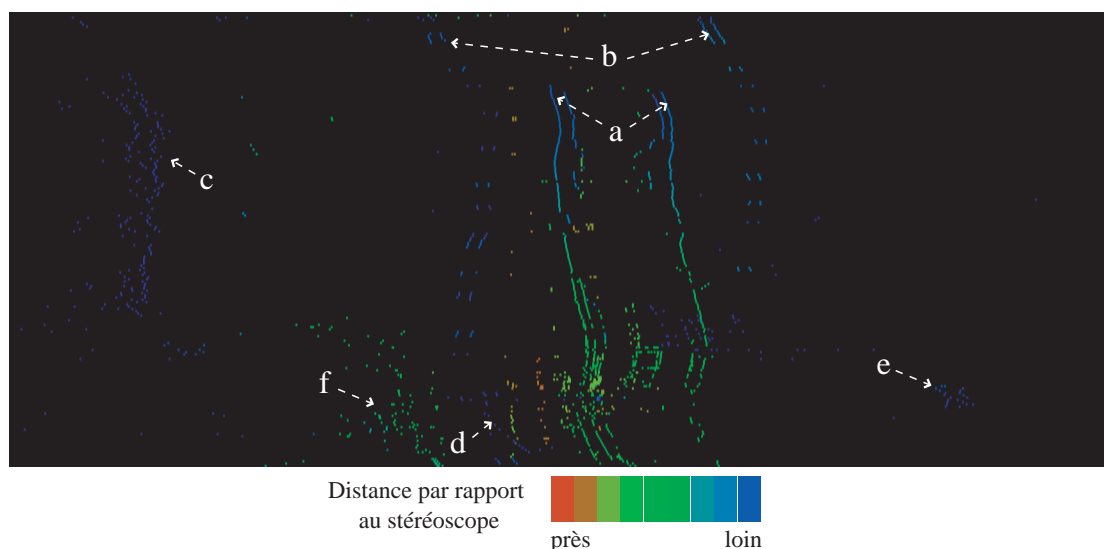


FIG. 3.38 : Résultat de la mise en correspondance de la séquence stéréoscopique "Approche d'un véhicule" obtenu grâce au codage binaire avec 100 chromosomes et 300 générations.

La figure 3.38 montre que les points de contour objet du véhicule (A) sont reconstruits correctement. En effet, leurs couleurs varient du bleu au vert clair, indiquant son rapprochement au cours du temps du véhicule prototype (a). D'autre part, les deux bandes en pointillés qui séparent les voies de l'autoroute sont également bien reconstruites (b). La ligne qui délimite le terre-plein central est également détectée, du moins lorsqu'elle est située dans le champ commun aux deux caméras (c). Le véhicule qui se rabat devant le véhicule (A) est visible en bas de l'image, sous la forme d'une ligne continue de couleur bleue correspondant à son extrémité gauche (d). L'extrémité de sa longue ombre est visible en bas, à l'extrême droite de l'image (e). Le véhicule qui double le véhicule prototype est visible, en bas à gauche de l'image (f).

Comme pour la séquence "Piéton", afin de comparer les résultats de mise en correspondance de couples d'images-ligne de la séquence "Approche d'un véhicule" pour différentes valeurs de $Taille_{pop}$ et de N_{gen} , on indique les meilleures valeurs atteintes par la fonction d'évaluation pour chacun de ces couples (cf. figures 3.40 et 3.41). Les courbes obtenues sont pratiquement confondues, indiquant que les meilleures valeurs atteintes par la fonction d'évaluation sont pratiquement les mêmes pour la majeure partie des couples, et ceci pour les différentes valeurs de $Taille_{pop}$ et de N_{gen} . Cette constatation est particulièrement vraie pour les valeurs correspondant aux couples compris entre les rangs 1 et 135 ainsi qu'aux couples compris entre les rangs 181 et 200. Pour les autres couples, on remarque que, globalement, l'augmentation de

la taille de la population ou/et du nombre de générations permet d'améliorer les solutions du problème traité. Par exemple, pour le couple d'images-ligne de rang 164 contenant 20 points de contour dans chacune des images-ligne gauche et droite (cf. figure 3.39), la valeur de la fonction d'évaluation est passée de 235 pour $Taille_{pop} = 50$ et $N_{gen} = 300$ à environ 150 pour les autres valeurs de $Taille_{pop}$ et de N_{gen} (cf. figures 3.40 et 3.41). Cette amélioration concerne également le couple d'images-ligne de rang 161 contenant 22 (resp. 16) points de contour dans l'image-ligne gauche (resp. droite) car la valeur de la fonction d'évaluation est passée de 95 pour $Taille_{pop} = 50$ et $N_{gen} = 300$ à 3 environ pour $Taille_{pop} = 100$ et $N_{gen} = 1000$ (cf. figures 3.40 et 3.41).

Si nous adoptons le raisonnement utilisé lors du traitement de la séquence "Piéton" qui consiste à relever les valeurs minimale et maximale de chacune des courbes des figures 3.40 et 3.41, nous constatons que celles-ci sont égales approximativement à :

- 235 et -30 pour $Taille_{pop} = 50$ et $N_{gen} = 300$,
- 150 et -30 pour $Taille_{pop} = 50$ et $N_{gen} = 1000$,
- 150 et -30 pour $Taille_{pop} = 100$ et $N_{gen} = 300$,
- 150 et -35 pour $Taille_{pop} = 100$ et $N_{gen} = 1000$.

En d'autres termes, la courbe ayant les valeurs maximale et minimale les plus petites est celle qui correspond à $Taille_{pop} = 100$ et $N_{gen} = 1000$.

Le tableau de la figure 3.42 indique le temps nécessaire pour obtenir le résultat de mise en correspondance des primitives extraites de cette séquence composée de 200 images-ligne stéréoscopiques et ceci pour différentes valeurs de $Taille_{pop}$ et de N_{gen} . Ce tableau montre que l'augmentation de la taille de la population et du nombre de générations implique une augmentation considérable du temps de traitement. En effet, pour $Taille_{pop} = 50$ et $N_{gen} = 300$ le temps du traitement est de l'ordre de 415 secondes (c'est-à-dire 2 secondes par couple d'images-ligne) et il passe à environ 1793 secondes (c'est-à-dire 9 secondes par couple d'images-ligne) pour $Taille_{pop} = 100$ et $N_{gen} = 1000$. Ces valeurs correspondent à des traitements effectués avec un PC Intel-Pentium IV-2GHz.

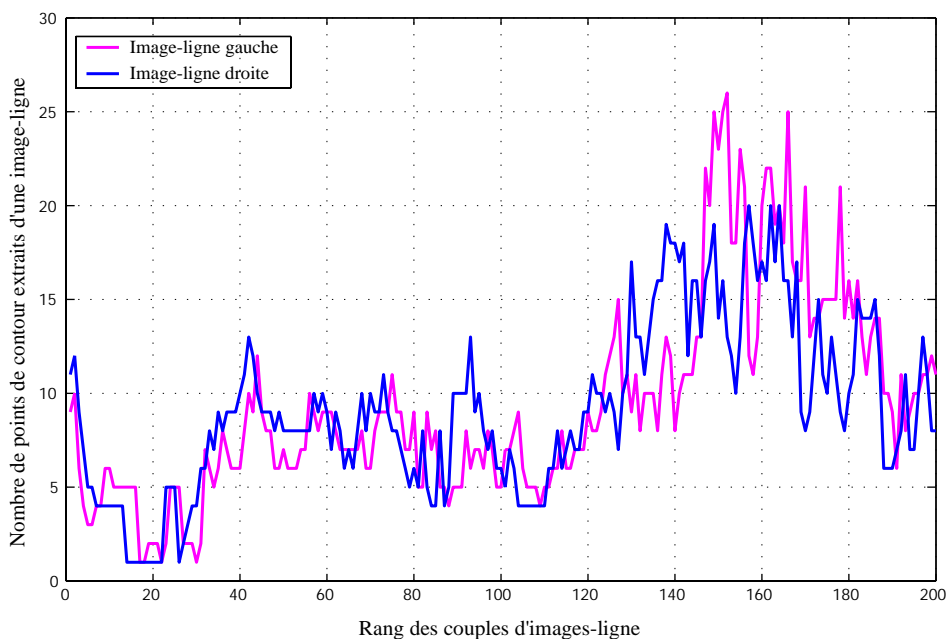


FIG. 3.39 : Nombre de points de contour extraits d'images-ligne de la séquence "Approche d'un véhicule".

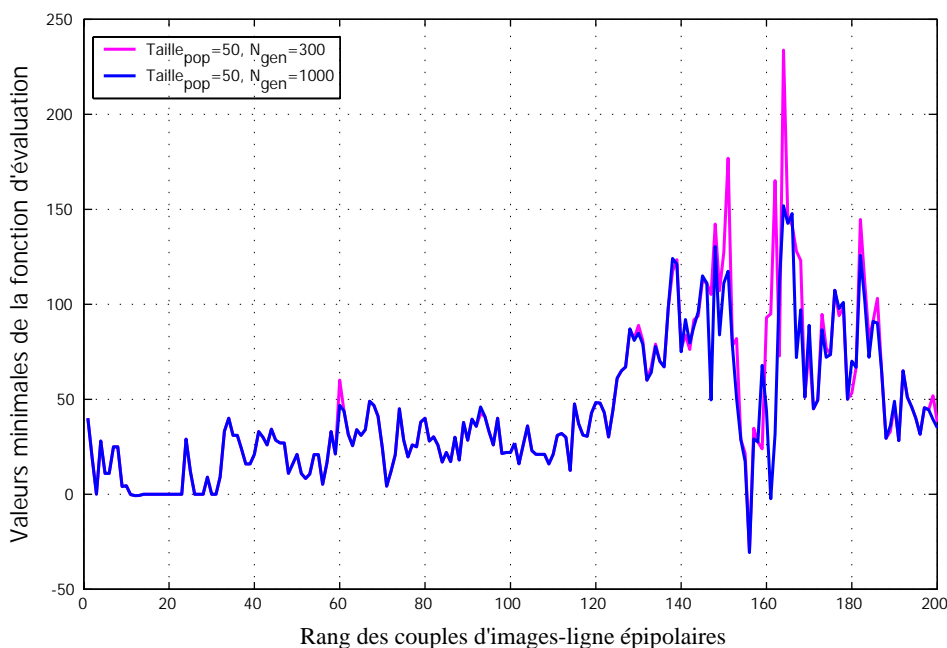


FIG. 3.40 : Valeurs minimales atteintes de la fonction d'évaluation pour chacun des couples d'images-ligne de la séquence "Approche d'un véhicule" avec d'une part $Taille_{pop} = 50$ et $N_{gen} = 300$ et d'autre part $Taille_{pop} = 50$ et $N_{gen} = 1000$.

3.9 Conclusion

Dans ce chapitre, nous avons présenté une nouvelle approche par algorithme génétique pour résoudre le problème de mise en correspondance des primitives extraites de couples d'images-

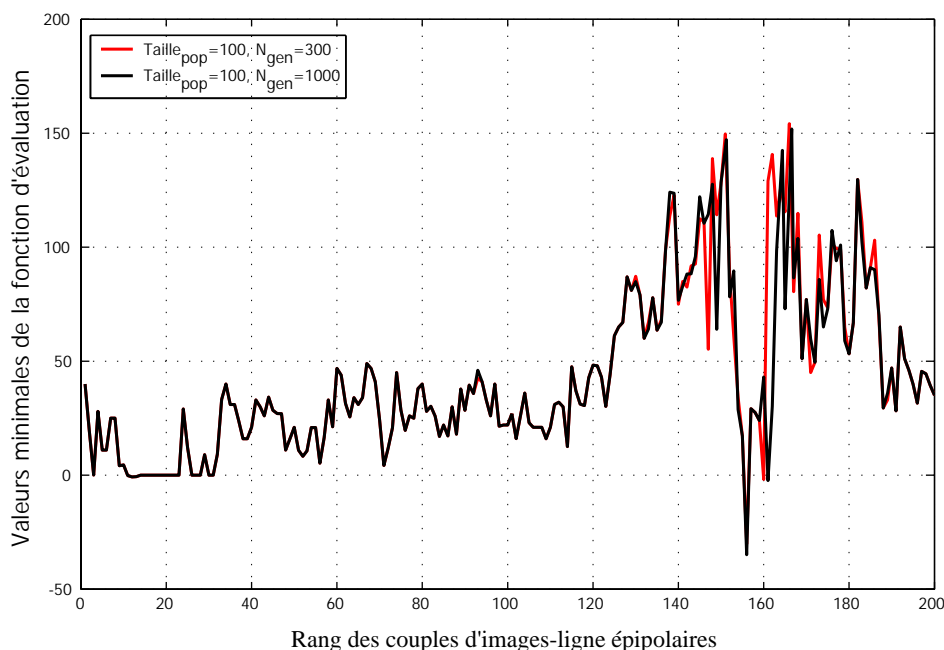


FIG. 3.41 : Valeur minimales atteintes de la fonction d'évaluation pour chacun des couples d'images-ligne de la séquence "Approche d'un véhicule" avec d'une part $Taille_{pop} = 100$ et $N_{gen} = 300$ et d'autre part $Taille_{pop} = 100$ et $N_{gen} = 1000$.

Temps de traitement [sec]		N_{gen}	
		300	1000
Taille _{pop}	50	415	952
	100	776	1793

FIG. 3.42 : Temps de traitement de la séquence "Approche d'un véhicule" pour différentes valeurs de $Taille_{pop}$ et de N_{gen} .

ligne épipolaires. Cette approche considère l'appariement entre deux images comme un problème d'optimisation où une fonction d'évaluation, intégrant des contraintes stéréoscopiques, doit être minimisée. Pour ce faire, il a tout d'abord fallu trouver un codage approprié permettant de représenter les solutions du problème sous forme de chromosomes. Ensuite, nous avons montré comment la prise en compte des contraintes locales de position et de signe de la dérivée permet de ne coder que des chromosomes qui représentent des solutions compatibles vis-a-vis de celles-ci. Ainsi, les chromosomes formés permettent de mieux explorer l'espace de recherche et l'algorithme converge ainsi vers une solution plus satisfaisante.

Après avoir établi le codage génétique du problème de mise en correspondance, nous avons

défini une fonction d'évaluation permettant de mesurer l'adaptation de chaque chromosome par rapport à l'objectif recherché, à savoir minimiser les erreurs commises en mettant en correspondance les primitives extraites de deux images-ligne épipolaires. Ensuite, nous avons décrit l'algorithme utilisé ainsi que les différents opérateurs génétiques nécessaires à son déroulement.

Les résultats de mise en correspondance obtenus lors des essais effectués sur deux séquences d'images-ligne stéréoscopiques réelles acquises grâce au stéréoscope linéaire développé au laboratoire I^3D sont encourageants. En effet, l'approche génétique développée fournit de résultats satisfaisants notamment quand il s'agit d'images-ligne contenant un nombre limité de points de contour. Mais, quand ce nombre augmente, la taille des chromosomes générés avec le codage binaire devient importante, ce qui nécessite une augmentation de la taille de la population et du nombre de générations afin que l'algorithme génétique puisse converger vers une solution acceptable. Cette augmentation des valeurs des deux paramètres se traduit par une augmentation considérable du temps de traitement nécessaire pour la phase de mise en correspondance. Pour pallier cette limitation qui vient du fait que l'algorithme génétique manipule des chromosomes binaires de tailles souvent importantes, nous proposons, dans le chapitre suivant, un autre codage que nous appellerons "codage entier".

Chapitre 4

Nouveau codage génétique pour l'appariement d'images stéréoscopiques

4.1 Introduction

Dans le chapitre précédent, nous avons présenté une nouvelle approche par algorithme génétique pour résoudre le problème de mise en correspondance en vision stéréoscopique. Ce problème est ramené à un problème d'optimisation où une fonction d'évaluation intégrant des contraintes stéréoscopiques doit être minimisée.

Nous avons proposé un codage permettant de représenter une solution au problème de mise en correspondance sous forme de chromosomes binaires. Les tailles de ces chromosomes dépendent du nombre de gènes qui les composent. Du fait qu'un gène binaire code une possibilité de mise en correspondance entre un point de contour extrait de l'image-ligne épipolaire gauche et un point de contour extrait de l'image-ligne épipolaire droite, la taille d'un chromosome binaire dépend du nombre de points candidats à l'appariement. Ainsi, la mise en correspondance d'un couple d'images-ligne épipolaires contenant un nombre important de points de contour nécessite la manipulation de chromosomes de grandes tailles.

Étant donné que la convergence d'un algorithme génétique vers une solution satisfaisante est sensible à qualité des chromosomes de la population initiale, nous avons utilisé les contraintes locales lors de la phase d'initialisation afin de créer une population de chromosomes composée uniquement de gènes respectant ces contraintes, ce qui a permis de réduire la taille des chromosomes binaires et d'explorer plus efficacement l'espace de recherche de solutions.

Le codage binaire a été employé pour appairer les couples d'images-ligne épipolaires de la séquence "Piéton" ainsi que ceux de la séquence "Approche d'un véhicule". Les résultats de la mise en correspondance obtenus sont satisfaisants mais le temps de traitement est élevé. Pour

pallier cette limitation due à la nature binaire de chromosomes, nous proposons un nouveau codage permettant d'explorer plus efficacement l'espace de recherche des solutions en produisant des chromosomes entiers représentant, eux aussi, des solutions respectant les contraintes locales.

4.2 Nouveau codage génétique

Soient L_G et L_D deux listes de points de contour détectés dans les lignes épipolaires gauche et droite, respectivement. Nous rappelons que chaque élément de ces listes est un triplet (x, s_x, d_x) où x désigne la position du point de contour dans la ligne épipolaire correspondant. s_x , qui prend les valeurs $+1$ ou -1 , représente le signe de la dérivée en ce point et d_x est l'amplitude de la dérivée en ce point. Les éléments de chacune de ces deux listes sont ordonnés de manière croissante selon les positions des points de contour extraits de la ligne épipolaire correspondant. Le point de contour ayant la position la plus petite est placé en tête de la liste et ainsi de suite.

Soit G (resp. D) la liste des indices des points de contour de la liste L_G (resp. L_D) et soient N_G et N_D les nombres de points de contour dans les listes L_G et L_D , respectivement.

Avec le nouveau codage génétique, un chromosome codant une solution possible au problème de mise en correspondance n'est plus une chaîne binaire. En effet, le nouveau chromosome C , appelé "chromosome entier", est désormais composé de gènes prenant des valeurs entières. La taille N_{max} du chromosome C , qui dépend du nombre de gènes qui le composent, est définie comme : $N_{max} = \text{Max}(N_G, N_D)$ (cf. figure 4.1). La valeur du gène $i \in T_{max}$ du chromosome C où T_{max} est la liste ordonnée de manière croissante des valeurs entières de l'intervalle $[1, N_{max}]$, est une valeur entière de l'intervalle $[0, N_{min}]$ où $N_{min} = \text{Min}(N_G, N_D)$.

Soit T_{min} une liste ordonnée de manière croissante des valeurs entières de l'intervalle $[1, N_{min}]$, alors $C_i \in 0 \cup T_{min}$. En d'autres termes, T_{max} (resp. T_{min}) peut-être vu comme étant la liste des indices des points de contour de la liste la plus longue entre L_G et L_D (resp. la plus petite).

L'interprétation de ce nouveau chromosome est la suivante : si la valeur d'un gène i est nulle ($C_i = 0$), cela signifie que le $i^{\text{ème}}$ point de contour de la liste L_G (ou L_D) n'a aucun correspondant dans la liste L_D (ou L_G). Sinon, ce $i^{\text{ème}}$ point de contour de la liste L_G (ou L_D) et le $C_i^{\text{ème}}$ point de contour de la liste L_D (ou L_G) sont appariés.

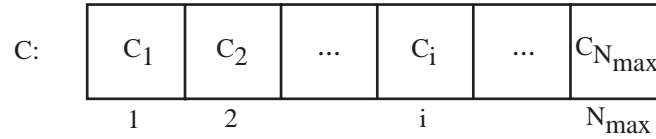


FIG. 4.1 : Chromosome entier constitué de N_{\max} gènes. La valeur du gène i , $i \in T_{\max} = \{1, 2, \dots, N_{\max}\}$ est une valeur entière C_i , $C_i \in \{0 \cup T_{\min}\}$ avec $T_{\min} = \{1, \dots, N_{\min}\}$.

Nous signalons qu'un chromosome entier, comme un chromosome binaire, code uniquement les appariements possibles qui respectent les contraintes locales. La figure 4.2 est un exemple d'un chromosome entier représentant des appariements possibles des points de contour de la figure 3.12(a). Dans cet exemple, $N_G = 6$ et $N_D = 5$. Ainsi, la valeur du gène i , $i \in T_{\max} = \{1, 2, 3, 4, 5, 6\}$ est C_i , $C_i \in \{0 \cup T_{\min}\}$ avec $T_{\min} = \{1, 2, 3, 4, 5\}$.

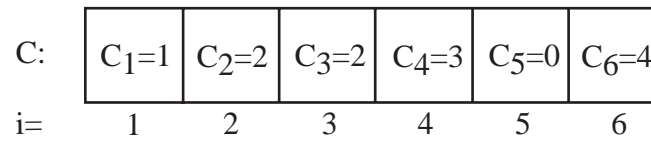
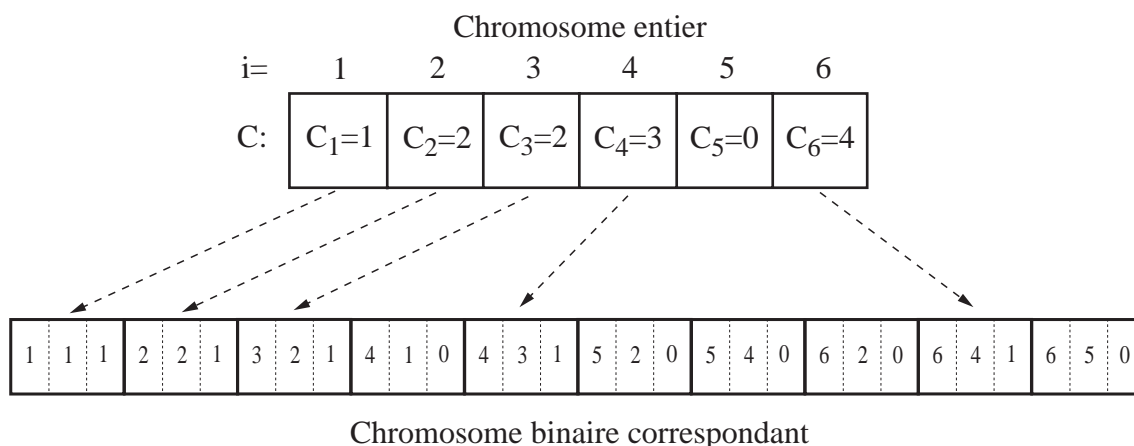


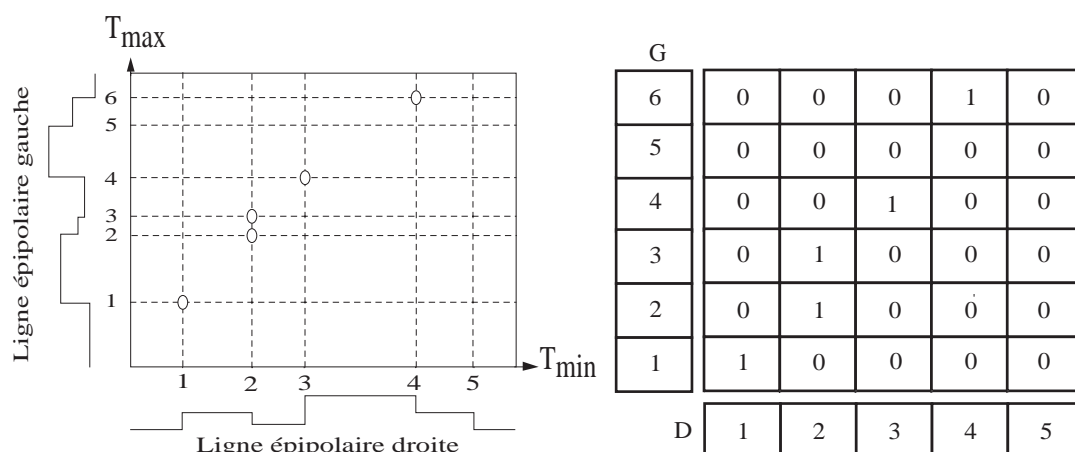
FIG. 4.2 : Exemple d'un chromosome généré grâce au codage entier où $i \in T_{\max} = \{1, 2, 3, 4, 5, 6\}$ et $C_i \in \{0 \cup T_{\min}\}$ avec $T_{\min} = \{1, 2, 3, 4, 5\}$.

On peut noter qu'il est facile de retrouver à partir d'un chromosome entier C le tableau binaire de mise en correspondance E . En effet, pour $i \in T_{\max}$ et $j \in T_{\min}$, $E_{ij} = 1$ si $C_i = j$, sinon $E_{ij} = 0$ (cf. figure 4.3). Il est important de signaler qu'un tableau de mise en correspondance (cf. figure 4.3(c)) obtenu à partir d'un chromosome entier présente moins d'appariements ambigus qu'un tableau binaire obtenu à partir d'un chromosome binaire (cf. figure 3.12(a)). En effet, on peut remarquer l'absence d'ambiguïtés sur les lignes du tableau de mise en correspondance obtenu grâce au codage entier (cf. figure 4.3(c)). De ce fait, un algorithme génétique manipulant des chromosomes entiers explore plus efficacement l'espace de recherche. Ainsi, la convergence vers une "bonne" solution est plus rapide.

Avant d'utiliser ce nouveau codage pour mettre en correspondance les couples d'images-ligne stéréoscopiques de la séquence "Piéton", nous allons adapter la fonction d'évaluation utilisée dans le cas du codage binaire afin qu'elle permette de mieux tenir compte de la modification apportée aux chromosomes.



(a) Chromosome entier et le chromosome binaire équivalent.



(b) Schéma illustrant les points de contour mis en correspondance. (c) Tableau binaire de mise en correspondance.

FIG. 4.3 : Exemple illustrant : (a) un chromosome entier et le chromosome binaire équivalent, (b) les points de contour appariés parmi ceux extraits d'un couple de lignes épipolaires gauche et droite avec $N_G = 6$ et $N_D = 5$ et (c) le tableau binaire de mise en correspondance.

4.3 Exemple

Dans ce paragraphe, nous nous servons d'un exemple afin d'illustrer l'intérêt du codage entier par rapport au codage binaire proposé précédemment pour traiter le problème de mise en correspondance d'images-ligne stéréoscopiques. La figure 4.4 montre le schéma de mise en correspondance d'un couple d'images-ligne stéréoscopiques. Sur ce schéma, nous montrons uniquement les possibilités d'appariement entre les points de contour des images-ligne gauche et droite respectant les contraintes locales. Les images-ligne stéréoscopique de cet exemple contiennent 15 points de contour pour l'image-ligne gauche et 17 points de contour pour l'image-ligne droite.

Pour cet exemple, un algorithme génétique de mise en correspondance mettant en œuvre le codage binaire développé précédemment, génère des chromosomes composés de seulement 81 gènes respectant les contraintes locales alors qu'on en dénombre 255¹ dans le cas où ces dernières ne sont utilisées. Mais un algorithme génétique utilisant le codage entier permet de réduire davantage la taille des chromosomes. En effet, le nombre de gènes nécessaires pour coder une solution au problème d'appariement des images-ligne de cet exemple est réduit à 17 et $N_{max} = N_D = 17$. Cette diminution importante de la taille des chromosomes (17 à la place de 81) permet de réduire considérablement le temps de traitement nécessaire pour la phase de mise en correspondance.

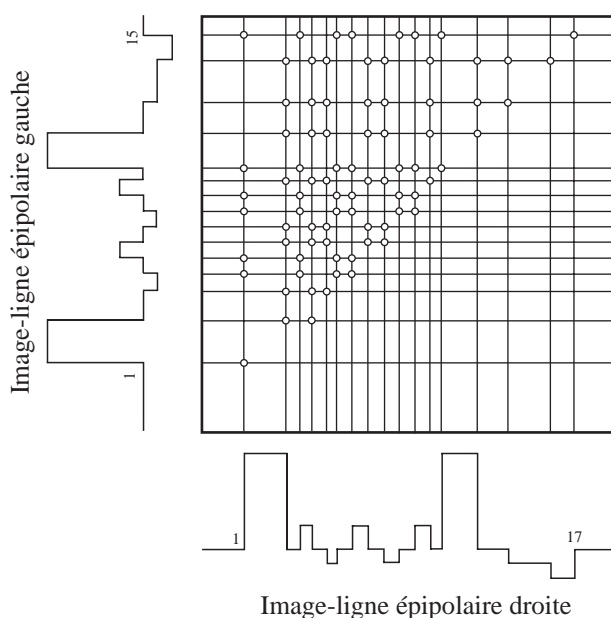


FIG. 4.4 : Exemple d'un schéma de mise en correspondance d'un couple d'images-ligne épipolaires avec $N_G = 15$ et $N_D = 17$.

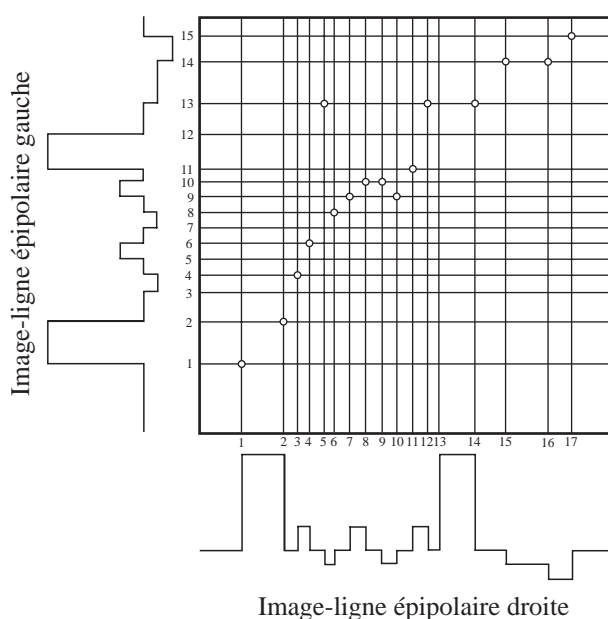
En plus de réduire le temps de traitement, les chromosomes entiers explorent plus efficacement l'espace de recherche de solutions grâce à l'élimination de cet espace de l'ensemble de solutions ne respectant pas la contrainte d'unicité selon les colonnes du schéma de mise en correspondance. Afin d'illustrer cette idée, un chromosome entier codant une solution possible au problème traité est montré (cf. figure 4.5(a)). Ce chromosome, composé de 17 gènes, code les possibilités d'appariement des points de contour extraits des images-ligne gauche et droite et qui respectent les contraintes locales. Dans ce chromosome, nous constatons que la valeur du gène numéro 10 est égale à 9. Ce gène apparie donc le point de contour d'indice 10 de la

¹15 × 17

liste des points de contour extraits de l'image-ligne droite ² avec le point de contour d'indice 9 de la liste des points de contour extraits de l'image-ligne gauche (cf. figure 4.5(a)). En d'autres termes, un point de contour de l'image-ligne droite ne peut être apparié qu'avec un point de contour de l'image-ligne gauche au maximum ³. Ainsi, le codage entier génère des chromosomes respectant implicitement la contrainte d'unicité selon les colonnes du schéma de mise en correspondance (cf. figure 4.5(b)). Grâce au codage entier, l'élimination des solutions ne respectant pas les contraintes d'unicité selon les colonnes permet de converger plus rapidement vers des solutions meilleures que celles obtenues avec le codage binaire et ceci pour un algorithme génétique utilisant la même taille des populations et le même nombre de générations.

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

(a) Chromosome entier.



(b) Schéma de mise en correspondance illustrant les points de contour appariés.

FIG. 4.5 : Exemple d'un chromosome entier codant une possibilité d'appariement d'un couple d'images-ligne épipolaires et le schéma de mise en correspondance équivalent.

Dans le paragraphe suivant, nous montrons que grâce à cette particularité du codage entier,

²Car c'est la liste la plus longue.

³Si la valeur du gène est de 0 alors le point de contour n'est apparié avec aucun point de contour de l'autre image-ligne.

nous pouvons améliorer le terme représentant la contrainte d'unicité dans la fonction d'évaluation. En effet, dans ce terme, la partie permettant de pénaliser les solutions ne respectant pas la contrainte d'unicité selon les colonnes du schéma de mise en correspondance peut être supprimée. Cette simplification permet d'accélérer encore une fois la procédure d'appariement d'images-ligne stéréoscopiques.

4.4 Adaptation de la fonction d'évaluation

Dans le paragraphe précédent, nous avons montré qu'il est facile de passer d'un chromosome entier à une représentation binaire sous forme de tableau. De ce fait, le calcul de la valeur de la fonction d'évaluation d'un chromosome entier C peut se faire à partir du tableau binaire de mise en correspondance équivalent.

Comme nous l'avons déjà signalé, les contraintes locales seront prises en compte lors de la phase d'initialisation de la population de chromosomes. Ainsi, la fonction d'évaluation sera composée uniquement des trois termes représentant les contraintes d'unicité, d'ordre et de continuité ainsi que du terme favorisant un nombre maximal d'appariements.

Le terme représentant la contrainte d'unicité peut être simplifié. En effet, comme nous l'avons précisé dans le paragraphe précédant, les lignes du tableau de mise en correspondance obtenues à partir d'un chromosome entier présentent moins d'ambiguïtés que celles dans le cas du codage binaire (*cf.* figures 4.3 et 3.12(a)). De ce fait, nous pouvons éliminer du terme représentant la contrainte d'unicité la partie pénalisant les solutions qui ne respectent pas la contrainte d'unicité selon les lignes du tableau binaire de mise en correspondance d'images-ligne épipolaires. Ainsi, la fonction d'évaluation prendra la forme :

$$\begin{aligned}
H_{tot} = & K_{uni} \sum_{j \in T_{min}} \left(1 - \sum_{i \in T_{max} / (i,j) \in \Delta} E_{ij} \right)^2 + \\
& K_{ord} \sum_{(i,j) \in \Delta} \sum_{(l,c) \in \Delta} O_{ijlc} E_{ij} E_{lc} - \\
& K_{cont} \sum_{(i,j) \in \Delta} \sum_{(l,c) \in \Delta} S_{ijlc} E_{ij} E_{lc} + \\
& K_m \left(N_{min} - \sum_{(i,j) \in \Delta} E_{ij} \right)^2
\end{aligned} \tag{4.1}$$

où K_{uni} , K_{ord} , K_{cont} , et K_m sont des coefficients positifs. Δ représente l'ensemble des appariements possibles entre un point de contour de la liste T_{max} et un point de contour de la liste T_{min} , c'est-à-dire l'ensemble de toutes les paires (i, j) satisfaisant les contraintes de position et de signe de la dérivée : $\Delta = \{(i, j) \in T_{max} \times T_{min}, (i, j) \text{ satisfait les contraintes locales}\}$.

Le premier terme de la fonction d'évaluation représente la contrainte d'unicité. Ce terme tend vers un minimum quand la somme des éléments du tableau de mise en correspondance dans chaque colonne est égale à 1.

Le deuxième terme permet de respecter la contrainte d'ordre. O_{ijlc} indique si l'ordre entre les deux couples (i, j) et (l, c) est respecté. Il est évalué par la formule suivante :

$$O_{ijlc} = |\text{signe}(x_i - x_l) - \text{signe}(x_j - x_c)|$$

avec :

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

Le troisième terme de la fonction d'évaluation représente la contrainte de continuité de la disparité. S_{ijlc} représente une mesure de compatibilité entre les deux paires de points de contour (i, j) et (l, c) candidats à l'appariement. Nous rappelons qu'il est exprimé en utilisant la fonction non linéaire suivante (cf. figure 3.15) :

$$S_{ijlc} = S(X_{ijlc}) = \frac{2}{1 + e^{(X_{ijlc} - \theta)}} - 1$$

où X_{ijlc} est la valeur absolue de la différence des disparités entre les deux couples (i, j) et (l, c) :

$$X_{ijlc} = |(x_i - x_l) - (x_j - x_c)|.$$

Le dernier terme de la fonction d'évaluation est utilisé pour renforcer les chromosomes qui codent un nombre maximal d'appariements. Ce terme tend vers un minimum lorsque le nombre d'appariements est égal à N_{min} .

4.5 Mise en correspondance pour la séquence "Piéton" avec codage entier et fonction d'évaluation adaptée

Nous avons utilisé l'algorithme génétique présenté dans le paragraphe 3.7 et le nouveau codage afin de mettre en correspondance les points de contour extraits de la séquence "Piéton". Nous n'avons pas modifié les valeurs des paramètres génétiques utilisées précédemment et nous

avons conservé les valeurs $Taille_{pop} = 100$ et $N_{gen} = 300$ afin de pouvoir comparer le résultat de la mise en correspondance obtenu grâce à ce codage avec celui qui a été obtenu en utilisant le codage binaire.

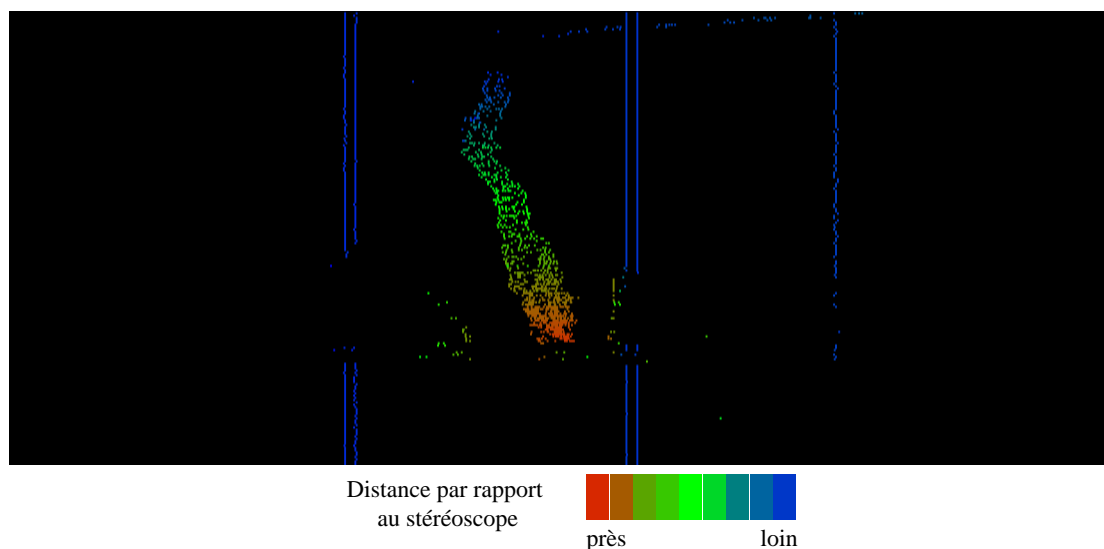


FIG. 4.6 : Résultat de la mise en correspondance de la séquence "Piéton" avec chromosomes entiers évalués à partir de leur équivalent binaire.

En comparant visuellement l'image de la mise en correspondance (*cf.* figure 3.25(a)) résultant de l'application de l'algorithme génétique utilisant le codage binaire avec celle de la figure 4.6 résultant du codage entier, nous constatons que cette dernière est meilleure. En effet, ce nouveau codage a permis d'apparier les images-ligne stéréoscopiques de la séquence avec moins de mauvais appariements. Les points de contour des bandes blanches sont bien appariés et leurs distances par rapport au stéréoscope sont constantes. En effet, ces points de contour correspondent aux points situés à l'intersection du plan de visée avec la chaussée (*cf.* figure 3.17). De même que pour les bandes, les points de contour désignant la transition entre la chaussée et la zone d'ombre à droite de la figure sont bien appariés. En effet, la couleur qui leur est attribuée indique leur éloignement du stéréoscope. En ce qui concerne le piéton, son mouvement vers le stéréoscope est bien reconstruit. Les points de l'image résultat qui le représentent changent progressivement de couleur. Celle-ci varie du rouge lorsqu'il est loin du stéréoscope au bleu quand il en est près, en passant par le vert, ce qui correspond bien à son déplacement vers le stéréoscope.

Pour mieux illustrer l'intérêt de ce nouveau codage par rapport au codage binaire, nous allons comparer l'évolution de la valeur de la fonction d'évaluation du meilleur chromosome au

cours des 300 générations. Cette comparaison, qui n'est valable que pour un couple d'images-ligne épipolaires, est possible car les deux fonctions d'évaluation ont le même minimum global car elles sont composées de termes représentant les mêmes contraintes (cf. §3.6.4 et §4.4).

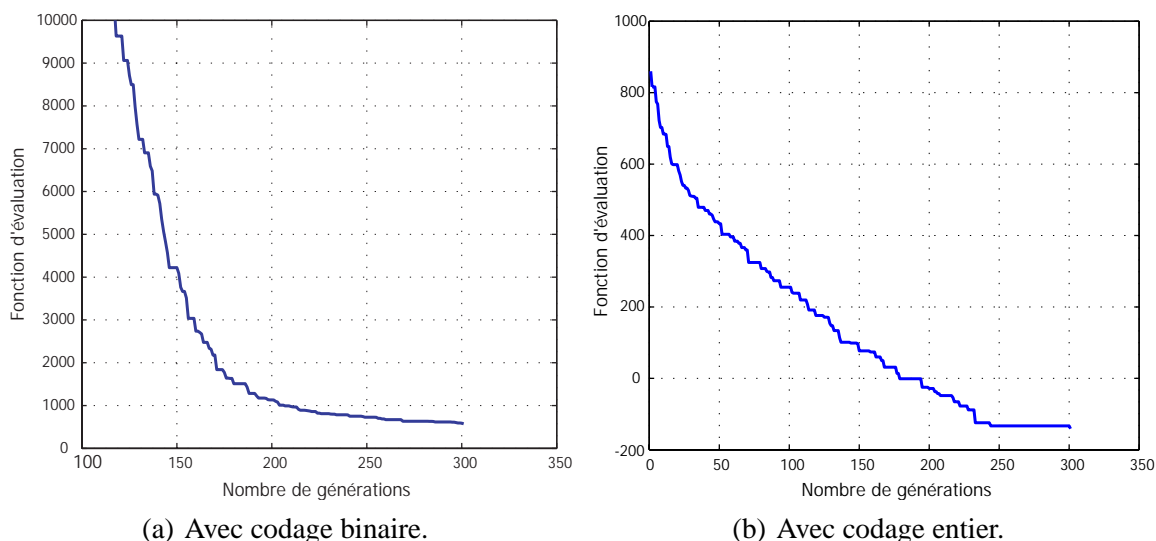


FIG. 4.7 : Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 100$ et $N_{gen} = 300$ pour le couple d'images-ligne de rang 134 de la séquence "Piéton".

En effet, au bout de 300 générations, la valeur de la fonction d'évaluation atteinte avec le codage binaire est d'environ 570 (cf. figure 4.7(a)), tandis qu'avec le codage entier cette valeur est de -140 (cf. figure 4.7(b)). Ce résultat signifie que ce nouveau codage entier a permis à l'algorithme génétique de mieux réussir la mise en correspondance du couple d'images-ligne épipolaires de rang 134 en atteignant une solution meilleure que celle obtenue avec le codage binaire. On remarque aussi en observant les courbes de cette figure que, d'une part, l'exploration de l'espace de recherche grâce au codage entier est beaucoup plus efficace et que d'autre part, la convergence vers une "bonne" solution est plus rapide (par exemple la valeur de la fonction d'évaluation au bout de 150 générations est de 70 avec le codage entier à la place de 4000 dans le cas du codage binaire).

4.5.1 Conclusion

Nous avons proposé un nouveau codage pour l'appariement des images stéréoscopiques qui permet de produire des chromosomes composés de gènes prenant des valeurs entières. Ces gènes codent plus astucieusement les appariements possibles des points de contour respectant les contraintes locales stéréoscopiques. Ainsi, les nouveaux chromosomes ont des tailles plus

petites et explorent de manière plus efficace l'espace de recherche des solutions.

Ce nouveau codage est utilisé pour mettre en correspondance les images-ligne stéréoscopiques de la séquence "Piéton". En comparant les résultats de la mise en correspondance obtenus grâce aux codages entier et binaire, nous avons constaté qu'avec le même nombre de générations et les mêmes tailles de populations, le nouveau codage a permis d'une part, de mieux explorer l'espace de recherche et d'autre part, de converger plus rapidement vers des solutions meilleures. En plus de l'amélioration du résultat de la mise en correspondance, ce nouveau codage a permis de réduire le temps de traitement d'un facteur de 1.5 par rapport à celui qui est nécessaire dans le cas du codage binaire (cf. figure 3.34). En effet, avec 100 chromosomes et 300 générations, le temps de traitement est de 1200 secondes environ (6 secondes par couple d'images-ligne à la place de 9 secondes) sur une machine équipée d'un micro-processeur Pentium IV-2GHz.

Malgré cette réduction du temps de traitement, ce dernier reste encore important. Pour tenter de le réduire, nous proposons, dans le paragraphe suivant, une nouvelle fonction d'évaluation qui permet de mesurer directement l'adaptation des chromosomes entiers à partir des gènes entiers qui les composent, plutôt que de passer par leurs équivalents binaires.

4.6 Nouvelle fonction d'évaluation des chromosomes entiers

La fonction d'évaluation est définie de telle sorte que ses minima correspondent à de bons appariements vis-à-vis des contraintes globales d'unicité, d'ordre et de continuité de la disparité. En utilisant le codage décrit dans le paragraphe 4.2, la nouvelle fonction d'évaluation peut-être exprimée sous la forme suivante :

$$\begin{aligned}
 H_{tot} = & K_{uni} \sum_{i=1}^{N_{min}} \left(1 - \sum_{j=1}^{N_{max}} U(i, C_j) \right)^2 + \\
 & K_{ord} \sum_{i=1}^{N_{max}-1} \sum_{j=i+1}^{N_{max}} O(C_i, C_j) - \\
 & K_{cont} \sum_{i=1}^{N_{max}-1} \sum_{j=i+1/C_i \neq C_j}^{N_{max}} S(C_i, C_j) + \\
 & K_m \left(N_{min} - \sum_{i=1}^{N_{max}} Z(C_i) \right)^2
 \end{aligned} \tag{4.2}$$

où K_{uni} , K_{ord} , K_{cont} et K_m sont des constantes de pondération positives.

Le premier terme de la fonction d'évaluation correspond à la contrainte d'unicité où $U(C_i, C_j)$ représente une pénalité quand cette contrainte n'est pas respectée :

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

Le deuxième terme permet de respecter la contrainte d'ordre. $O(C_i, C_j)$ indique si l'ordre des deux paires (i, C_i) et (j, C_j) est respecté :

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

Le troisième terme correspond à la contrainte de continuité de la disparité où $S(C_i, C_j)$ représente une mesure de compatibilité entre les deux paires (i, C_i) et (j, C_j) :

$$S(C_i, C_j) = S(X) = \frac{2}{1 + e^{(X-\theta)}} - 1,$$

Dans cette expression, X est la valeur absolue de la différence des disparités des paires (i, C_i) et (j, C_j) .

Le dernier terme de la fonction d'évaluation est utilisé pour favoriser les chromosomes qui codent un nombre maximal d'appariements. Ce terme tend vers un minimum lorsque le nombre d'appariements est égal à N_{min} . $Z(C_i)$ est défini comme suit :

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

Comme nous l'avons vu dans le chapitre précédent, le paramètre θ 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 bons appariements qui correspondent à des valeurs de X voisines de 0, et qu'une compatibilité minimale, c'est-à-dire proche de -1 , soit obtenue pour les mauvais appariements qui correspondent à des valeurs élevées de X (cf. figure 3.15).

4.7 Résultat obtenu sur la séquence "Piéton" grâce au codage entier et la nouvelle fonction d'évaluation

De nouveau, l'algorithme génétique décrit au paragraphe 3.7 est employé afin de mettre en correspondance les images stéréoscopiques de la séquence "Piéton". Cette fois-ci, pour mesurer

l'adaptation des chromosomes entiers, nous utilisons la nouvelle fonction d'évaluation définie dans le paragraphe précédent. La figure 4.8 représente la reconstruction 3D obtenue avec les mêmes valeurs des paramètres génétiques que celles utilisées précédemment (cf. § 3.8.1.2). Cette figure montre que le résultat de la reconstruction est très proche de celui que nous avons obtenu avec des chromosomes entiers évalués avec la fonction décrite dans le paragraphe 4.4. En effet, les points de contour des bandes sont bien appariés ainsi que ceux qui correspondent au piéton. Comme sur les résultats obtenus précédemment, on remarque la présence de mauvais appariements lorsque des occlusions se produisent, c'est-à-dire lorsque le piéton cache l'une des bandes blanches à l'une des caméras.

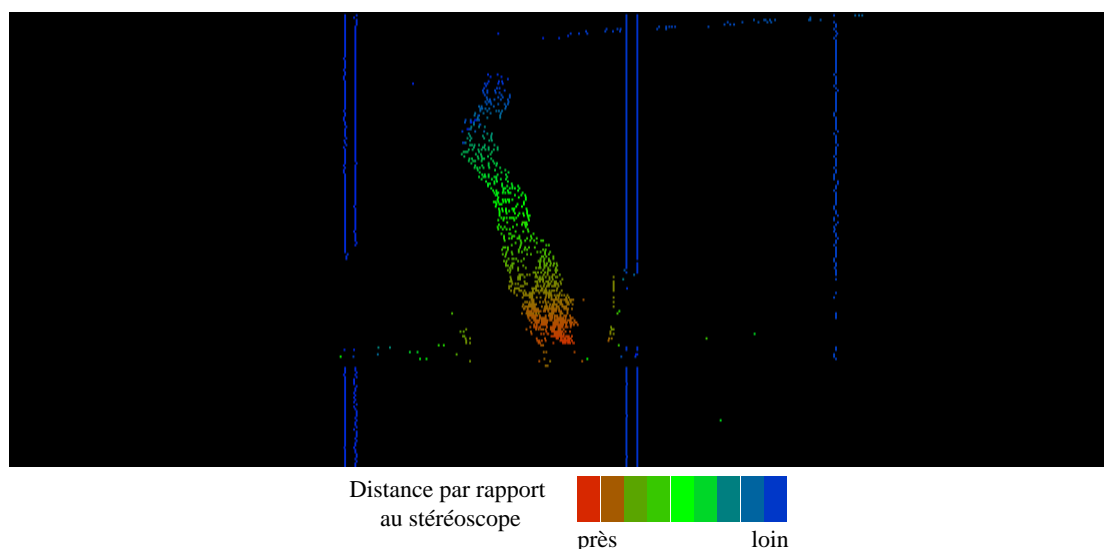


FIG. 4.8 : Résultat de la mise en correspondance de la séquence "Piéton" avec le codage entier et la nouvelle fonction d'évaluation.

Pour comparer les comportements de cette nouvelle fonction d'évaluation et de celle décrite dans le paragraphe 4.4, nous considérons les courbes des figures 4.9 et 4.7(b). Ces courbes représentent l'évolution des valeurs des deux fonctions d'évaluation du meilleur chromosome obtenues lors de l'appariement du couple d'images-ligne épipolaires de rang 134 de la séquence "Piéton". Ces courbes montrent premièrement que l'évolution des valeurs des deux fonctions est pratiquement la même et qu'au bout de 300 générations, les valeurs des deux fonctions se rapprochent de la valeur -150 .

Grâce à cette nouvelle fonction, l'évaluation de la qualité d'un chromosome entier est devenue possible sans avoir recours à sa représentation binaire. De plus, le temps de traitement est réduit d'un facteur de 15 par rapport à celui qui est nécessaire dans le cas du codage binaire

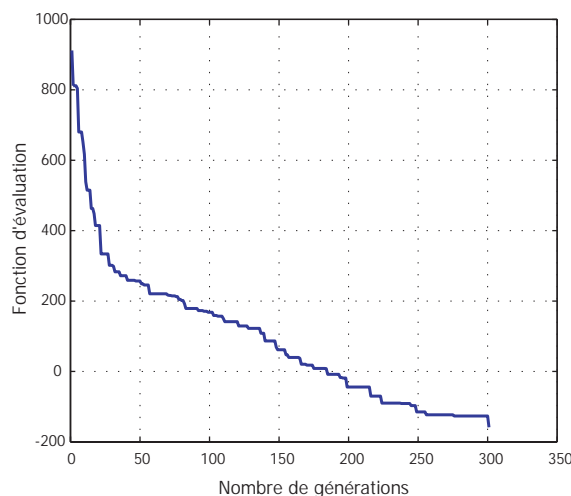


FIG. 4.9 : Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 100$ et $N_{gen} = 300$ pour le couple d'images-ligne de rang 134 de la séquence "Piéton".

(cf. figure 3.34). En effet, le temps de traitement est de 120 secondes environ, c'est-à-dire 0.6 seconde par couple d'images-ligne à la place de 9 secondes.

Dans le paragraphe suivant, nous allons utiliser le codage entier et la nouvelle fonction d'évaluation afin d'apparier les images-ligne stéréoscopiques de la séquence "Approche d'un véhicule".

4.8 Résultat obtenu sur la séquence "Approche d'un véhicule"

La séquence "Approche d'un véhicule" dont les images-ligne sont présentées dans le paragraphe 3.8.3 a déjà été traité par la procédure génétique de mise en correspondance en utilisant le codage binaire. Dans ce paragraphe, nous retraits cette séquence par la même procédure mais cette fois-ci avec le codage entier présenté dans le paragraphe 4.2.

La figure 4.10 montre le résultat de la mise en correspondance obtenu grâce au nouveau codage. Nous constatons que ce résultat est très proche de celui obtenu précédemment avec le codage binaire. En effet, on constate sur cette figure que les points de contour du véhicule (A) (cf. figure 3.36) sont bien appariés (a) car ils ont des couleurs cohérentes avec le rapprochement du véhicule prototype au cours du temps. Les deux bandes (b), en pointillés, qui séparent les voies de l'autoroute et la ligne qui délimite le terre-plein central (c) sont également bien appariées et leur détection est stable tout au long de la séquence. La couleur bleue qui leur est attribuée indique que ces bandes et la ligne sont éloignées du stéréoscope. Les points de contour

correspondant à l'extrémité gauche du véhicule (B) qui se rabat devant le véhicule (A) sont bien appariés (cf. figure 3.36). Ils sont visibles en bas de l'image avec une couleur bleue indiquant l'éloignement du véhicule (B) par rapport au véhicule prototype. Nous rappelons que l'extrémité droite du véhicule (B) n'est pas visible car le véhicule (A) la masque à la caméra droite. L'extrémité de la longue ombre du véhicule (B) est visible en bas, à l'extrême droite de l'image (e). Les points de contour du véhicule (C) qui double le véhicule prototype sont bien localisés et visibles en bas à gauche de l'image (f). Les couleurs qui leur sont attribuées sont parfaitement cohérentes avec leur disposition spatiale.

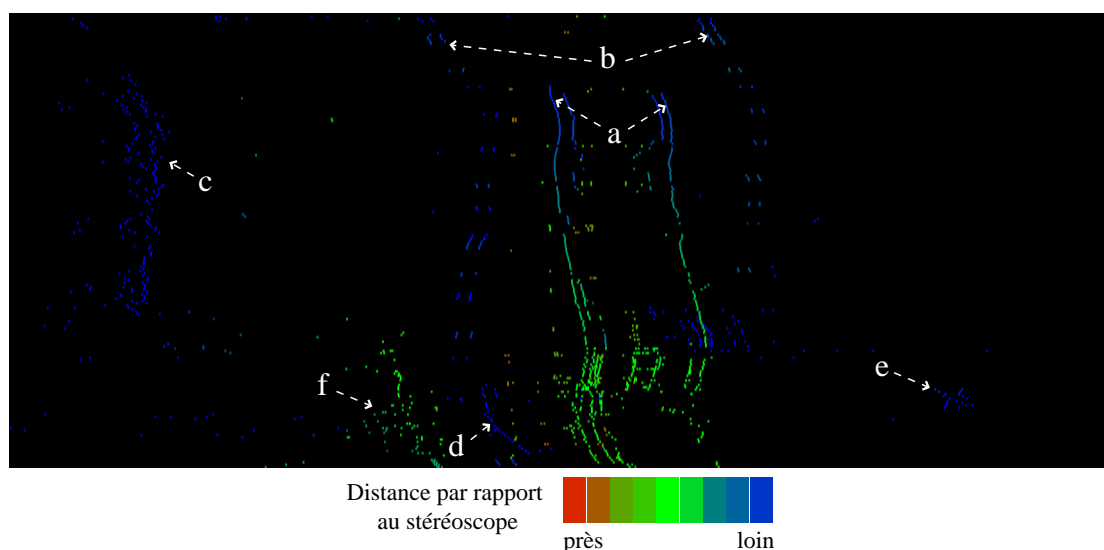


FIG. 4.10 : Résultat de la mise en correspondance de la séquence stéréoscopique "Approche d'un véhicule".

Grâce au codage entier et à la nouvelle fonction d'évaluation, le temps de traitement nécessaire pour la mise en correspondance des images-ligne pour la séquence "Approche d'un véhicule" est passé de 776 secondes à 90 secondes environ, c'est-à-dire 0.45 secondes par couple d'images-ligne à la place de 4 secondes.

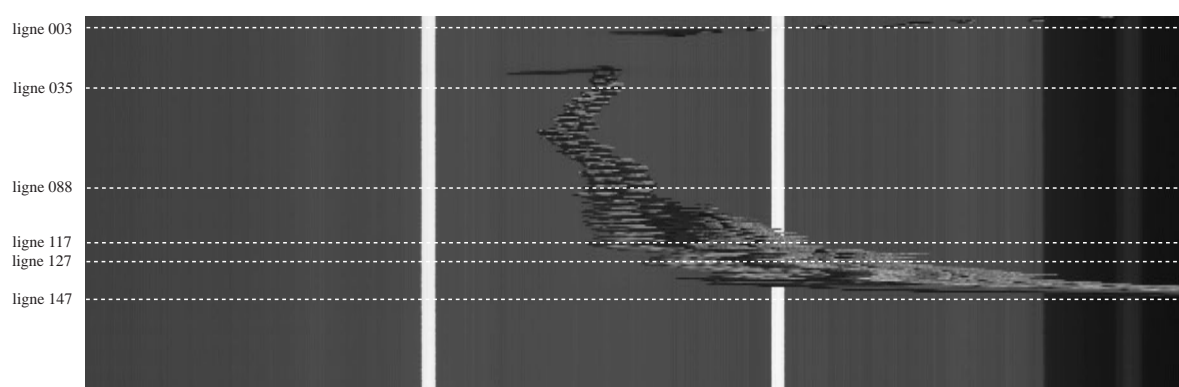
Dans le paragraphe suivant, nous allons discuter le réglage de deux paramètres génétiques qui sont la taille de la population et le nombre de générations.

4.9 Réglage des paramètres $Taille_{pop}$ et N_{gen}

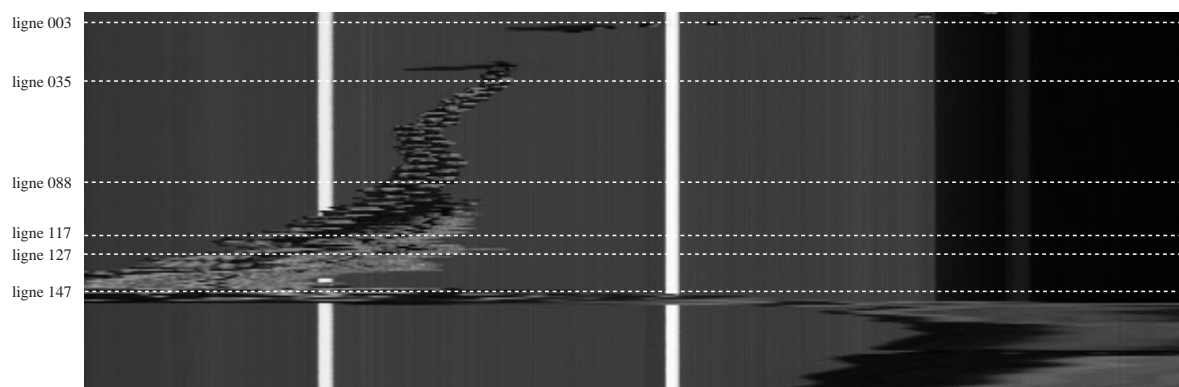
Lors du traitement de la séquence "Piéton", les valeurs des paramètres $Taille_{pop}$ et N_{gen} ont été fixés à des valeurs identiques pour les deux codages génétiques utilisés afin de rendre possible la comparaison des résultats de mise en correspondance ($Taille_{pop} = 100$ et $N_{gen} = 300$).

Cependant, ces valeurs ne sont pas forcément adaptées à tous les couples d'images stéréoscopiques que nous souhaitons traiter.

Pour expliquer le problème de l'ajustement de ces paramètres, nous allons analyser les courbes de la figure 4.12 illustrant l'évolution de la valeur de la fonction d'évaluation de quatre couples d'images-ligne extraits de la séquence "Piéton". Les couples choisis, de rangs 35, 117, 127 et 129 sont de complexité croissante en termes de nombre de points de contour extraits des images-ligne gauche (N_G) et droite (N_D) (cf. figures 4.11(a) et 4.11(b)).



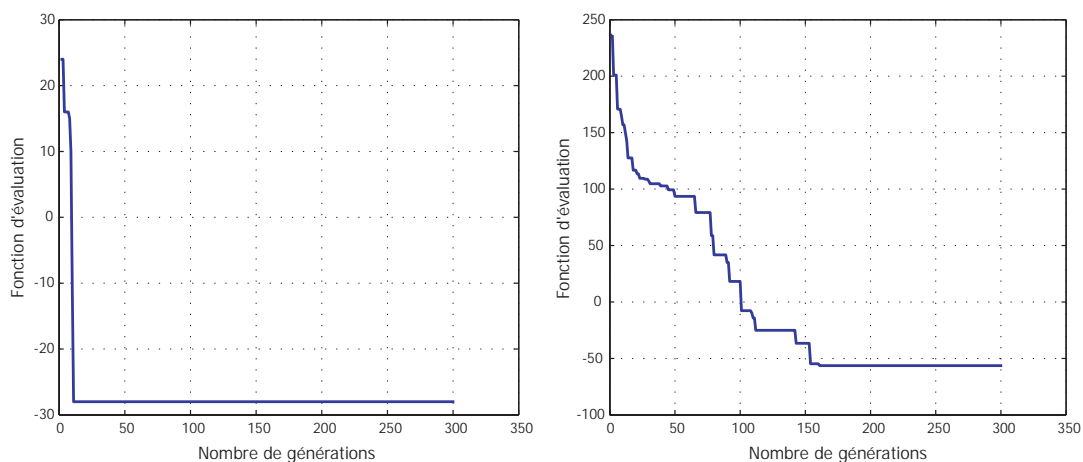
(a) Séquence gauche.



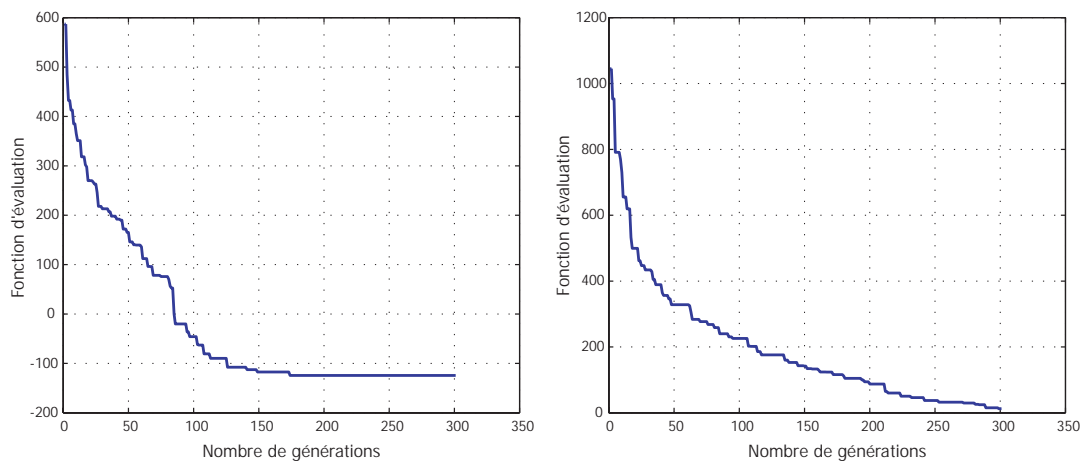
(b) Séquence droite.

FIG. 4.11 : Séquence de déplacement d'un piéton.

La figure 4.12(a) montre que pour le couple d'images de rang 35, 10 générations environ ont suffi pour que l'algorithme génétique converge vers une solution satisfaisante, tandis que pour les couples de rangs 117 et 127, il a fallu attendre environ 175 générations pour que la valeur de la fonction d'évaluation arrête d'évoluer (cf. figures 4.12(b) et 4.12(c)). En ce qui concerne le couple de rang 129, nous constatons que 300 générations ne sont peut-être pas suffisantes pour atteindre une "bonne" solution. En effet, la courbe de la figure 4.12(d) montre qu'au cours des 300 générations, la valeur de la fonction d'évaluation n'a pratiquement pas arrêté d'évoluer.



(a) Couple de rang 35 avec $(N_G = 8, N_D = 8)$. (b) Couple de rang 117 avec $(N_G = 17, N_D = 17)$.



(c) Couple de rang 127 avec $(N_G = 24, N_D = 24)$. (d) Couple de rang 129 avec $(N_G = 33, N_D = 24)$.

FIG. 4.12 : Évolution de la fonction d'évaluation du meilleur chromosome pour quatre couples d'images-ligne extraits de la séquence "Piéton".

En d'autres termes, les valeurs des deux paramètres génétiques pourraient être réglées en fonction du nombre de points de contour candidats à l'appariement. Pour des couples d'images où N_G et N_D sont petits, $Taille_{pop}$ et N_{gen} pourraient être petits car, pour ces images, la convergence vers une bonne solution est souvent rapide, ce qui n'est pas le cas quand il s'agit d'apparier des couples d'images contenant un nombre important de points de contour.

Avant de formuler les équations permettant de calculer les valeurs de ces paramètres, nous attirons l'attention du lecteur sur un cas particulier. En effet, lorsque la différence entre le nombre de points de contour extraits des images gauche et droite augmente, l'appariement de ces images devient plus compliqué. De ce fait, la procédure génétique de mise en correspon-

dance nécessite d'avantage de chromosomes et d'avantage de générations. Pour mieux illustrer ce cas de figure, nous choisissons quatre autres couples d'images de rangs 3, 88, 147 et 150 extraits de la séquence "Piéton". Pour les deux premiers couples, la différence entre le nombre de points de contour extraits des images gauche et droite est nulle, tandis que pour les deux autres couples, cette différence est de 6 points.

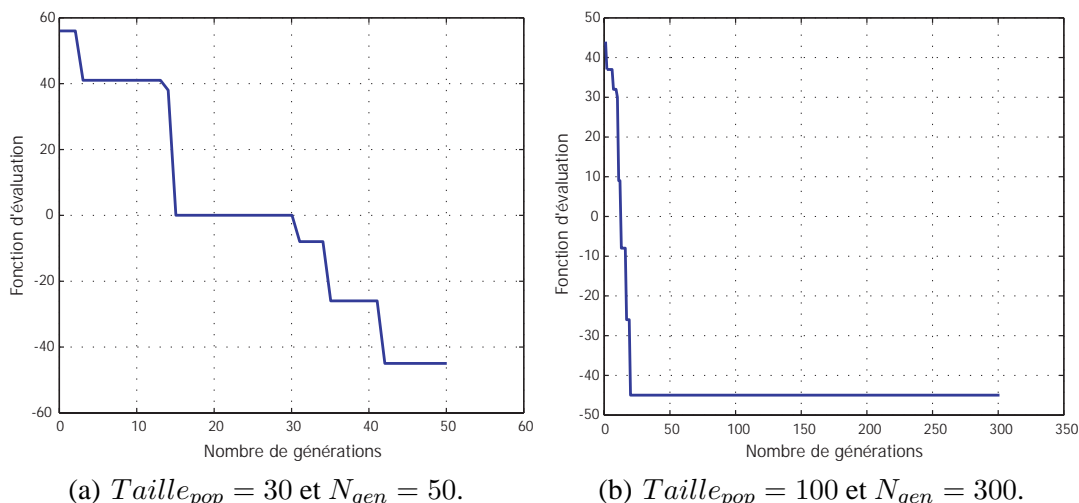


FIG. 4.13 : Évolution de la fonction d'évaluation du meilleur chromosome pour le couple d'images-ligne de rang 3 avec $N_G = 10$ et $N_D = 10$.

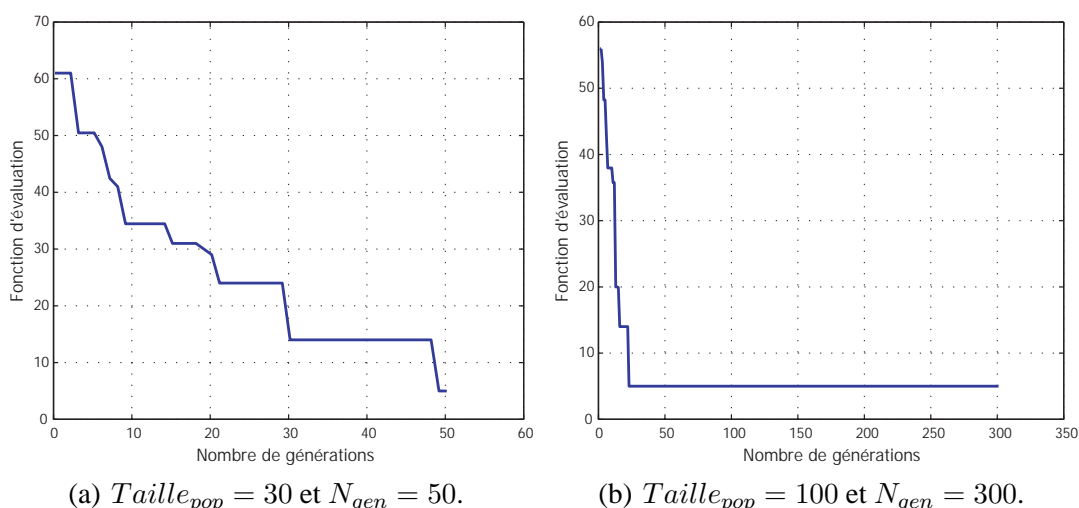


FIG. 4.14 : Évolution de la fonction d'évaluation du meilleur chromosome pour le couple d'images-ligne de rang 88 avec $N_G = 10$ et $N_D = 10$.

En comparant l'évolution de la valeur de la fonction d'évaluation du meilleur chromosome dans le cas où $Taille_{pop} = 30$ et $N_{gen} = 50$ et dans le cas où $Taille_{pop} = 100$ et $N_{gen} = 300$, nous constatons que pour les deux premiers couples, la meilleure valeur de la fonction

d'évaluation atteinte est la même dans les deux cas et ceci malgré la diminution de la taille des populations et du nombre des générations (cf. figures 4.13 et 4.14). Par contre, pour les deux autres couples d'images, nous constatons que dans le cas où la taille des populations et le nombre des générations sont réduits, l'algorithme n'a pas pu converger vers une "bonne" solution comme celle trouvée dans le cas où $Taille_{pop} = 100$ et $N_{gen} = 300$, et ceci malgré que N_G et N_D sont très proches, voir plus petits que pour les deux premiers couples d'images (cf. figures 4.15 et 4.16).

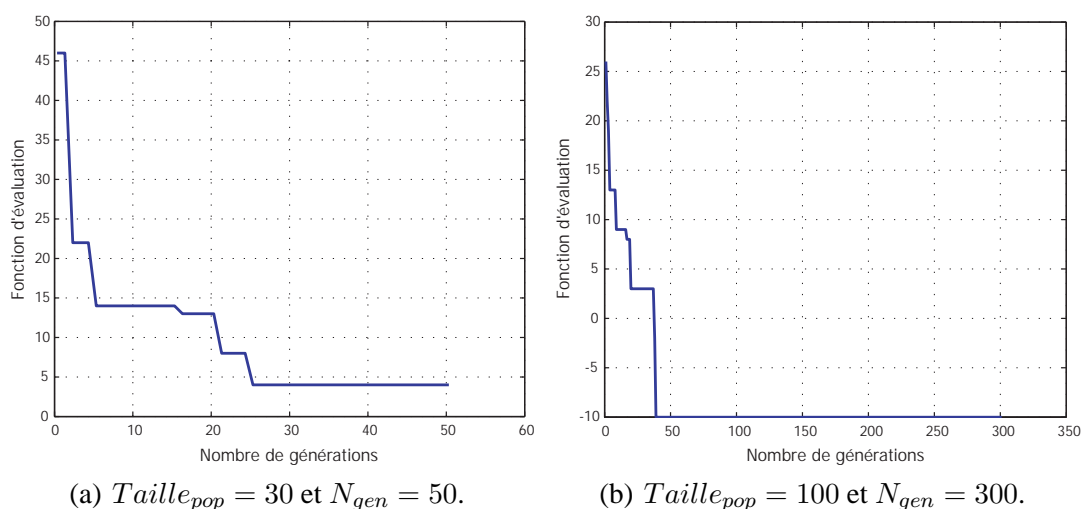


FIG. 4.15 : Évolution de la fonction d'évaluation du meilleur chromosome pour le couple d'images-ligne de rang 147 avec $N_G = 5$ et $N_D = 11$.

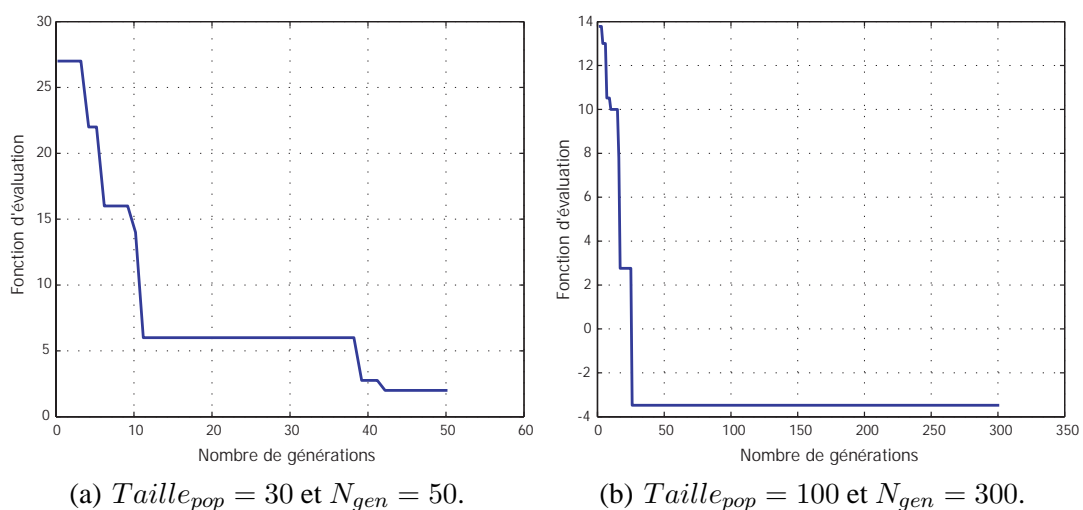


FIG. 4.16 : Évolution de la fonction d'évaluation du meilleur chromosome pour le couple d'images-ligne de rang 150 avec $N_G = 5$ et $N_D = 11$.

En tenant compte des éléments présentés ci-dessus, les deux formules empiriques que nous

proposons pour régler la taille des populations et le nombre des générations sont les suivantes :

$$Taille_{pop} = A \times (N_G + N_D) + B \times |N_G - N_D| \quad (4.3)$$

$$N_{gen} = C \times Taille_{pop} \quad (4.4)$$

où A, B et C sont des constantes positives. Après de nombreux tests effectués différents couples d'images stéréoscopiques, ces constantes ont été fixées comme suit :

- A=5 ;
- B=2 ;
- C=1.

En ce qui concerne les paramètres P_{crois} , P_{mut} et $Selection_{det}$, les valeurs retenues sont :

- $P_{crois} = 0.6$;
- $P_{mut} = 1/\text{taille du chromosome}$;
- $Selection_{det} = 0.10$.

Les résultats de l'application de ce réglage pour l'appariement des primitives extraites des séquences stéréoscopiques "Piéton" et "Approche d'un véhicule" sont présentés dans le chapitre suivant car ils sont comparés avec les résultats obtenus par une autre approche (cf. paragraphes 5.7.1 et 5.7.2).

Dans les paragraphes suivants, nous décrivons d'autres d'images stéréoscopiques réelles et de synthèse et nous présentons les résultats d'extraction des points de contour ainsi que les résultats d'appariement de ces primitives.

4.10 Autres résultats expérimentaux

La procédure de mise en correspondance utilisée pour traiter les images-ligne épipolaires des séquences "Piéton" et "Approche d'un véhicule" est appliquée pour appairer de nouveaux couples d'images stéréoscopiques générées en respectant la configuration particulière permettant d'éviter le calcul préalable des équations des droites épipolaires (cf. figure 2.2). Les deux premiers sont deux couples d'images stéréoscopiques de synthèses appelées successivement "Voie" et "Couloir". Ces couples d'images sont générés par un *programme de lancer de rayons*⁴.

⁴Ray tracing program.

Le troisième couple, est le couple d'images "Village", qui est constitué d'images prises avec une seule caméra qui peut être déplacée grâce à une plate-forme motorisée. Ensuite, la même procédure est utilisée pour mettre en correspondance des primitives extraites d'un couple d'images stéréoscopiques réelles appelé "Rails". Ces couples d'images stéréoscopiques sont de complexité croissante en termes de nombre de points d'intérêt candidats à l'appariement. Elles ont été sélectionnées afin de montrer la capacité de notre procédure génétique à fournir des résultats de mise en correspondance satisfaisants.

4.10.1 Traitement du couple d'images "Voie"

4.10.1.1 Description et extraction des points de contour

Le couple d'images de la figure 4.17 simule une voie de métro ou une route et des obstacles qu'un conducteur humain ou un véhicule automatisé peut rencontrer devant lui sur son trajet. Ce couple d'images de synthèse "Voie" représente des objets parallélépipédiques placés sur une voie et de part et d'autre de celle-ci, à des distances différentes par rapport à un stéréoscope fictif.

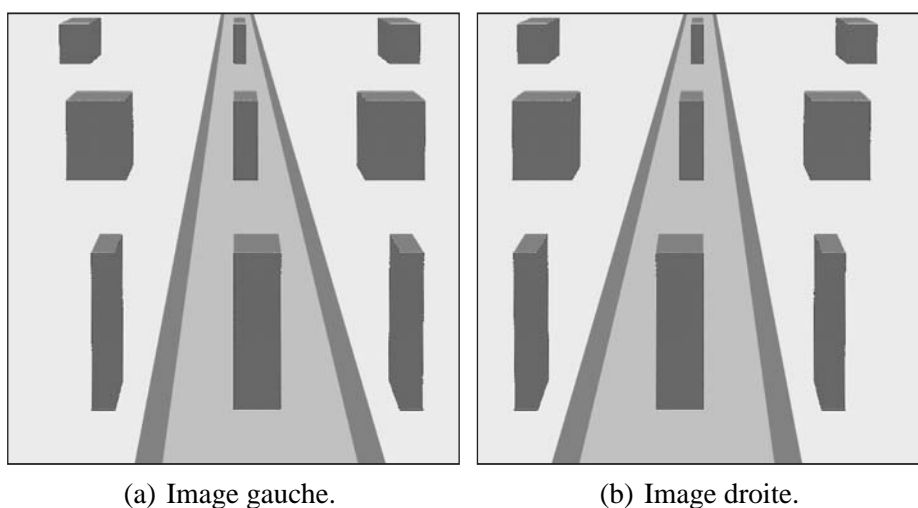
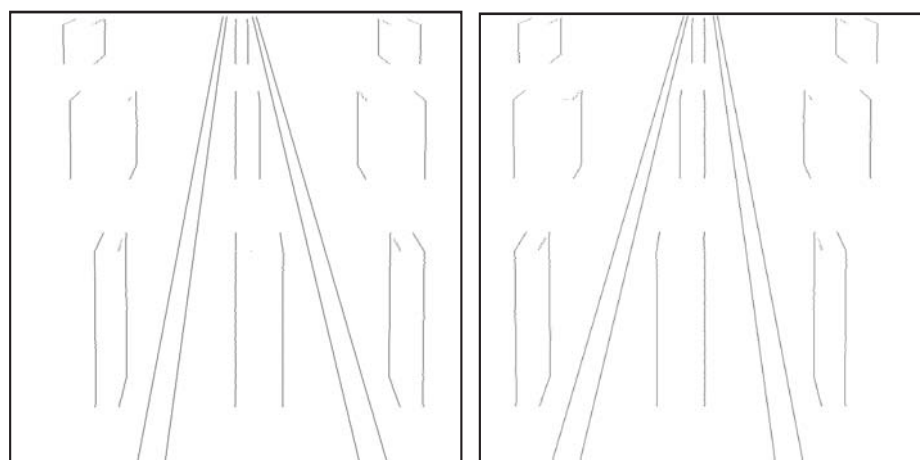


FIG. 4.17 : Couple d'images "Voie".

La figure 4.18 illustre les résultats du traitement obtenus sur ce couple d'images stéréoscopiques. Pour obtenir une bonne lisibilité, les points de contour extraits sont représentés par des points noirs quelque soit la valeur de l'amplitude de leur dérivée.

Le module d'extraction de primitives a permis une bonne détection des points de contour associés aussi bien aux différents objets présents dans la scène qu'à la voie située au milieu.



(a) Points de contour extraits de l'image gauche. (b) Points de contour extraits de l'image droite.

FIG. 4.18 : Extraction de contours du couple d'images "Voie".

L'absence des contours horizontaux sur la figure 4.18 est due au fait que la procédure d'extraction de contours traite les images matricielles ligne par ligne grâce au filtre de Deriche. En d'autres termes, pour chacune des lignes qui composent une image matricielle, la procédure de détection ne conserve que les points ayant une variation significative de niveau de gris par rapport à leurs voisins situés uniquement sur cette ligne. Ainsi, l'empilement du résultat de l'extraction de contour de chacune de ces lignes ne fait apparaître que les contours verticaux et obliques.

4.10.1.2 Résultat de la mise en correspondance

Le couple d'images "Voie" présenté dans le paragraphe précédent, est traité par la procédure génétique décrite au paragraphe 3.7. Pour soucis de clarté dans la représentation du résultat de la reconstruction 3D, la distance par rapport au stéréoscope est représentée en couleur. Celle-ci varie du rouge, qui correspond à la distance la plus petite par rapport au stéréoscope, au bleu qui correspond à la distance la plus grande.

La figure 4.19 montre que les points de contour correspondants aux différents objets parallélépipédiques ainsi qu'à la voie sont bien appariés et ceci dans toute la scène. On peut remarquer aussi que le nombre de mauvais appariements ainsi que le nombre d'appariements manqués est très limité.

Ce couple, composé de 600 d'images-ligne stéréoscopiques, est traité en 50 secondes, c'est-à-dire que le temps de traitement d'un couple d'images-ligne est de l'ordre de 0.08 seconde.

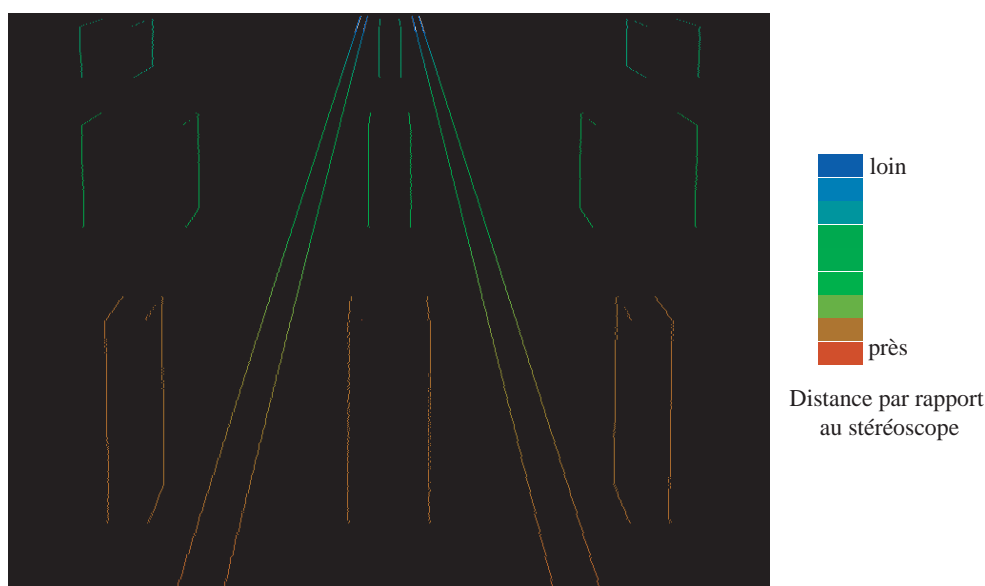


FIG. 4.19 : Résultat de la mise en correspondance du couple d'images "Voie".

4.10.2 Traitement du couple d'images "Couloir"

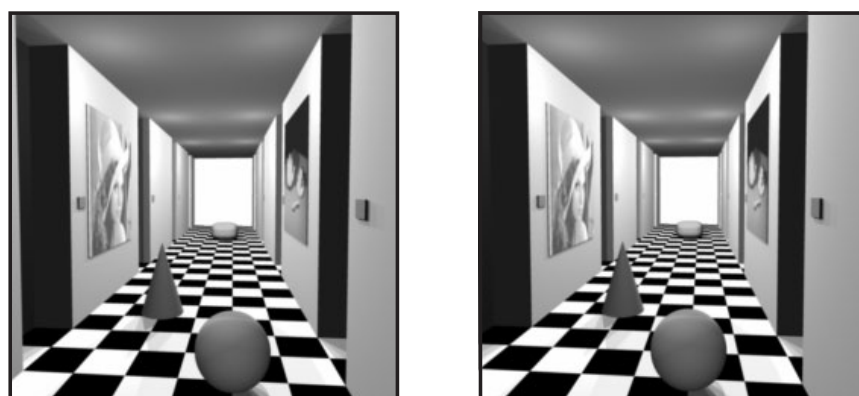
4.10.2.1 Description et extraction des points de contour

Le couple d'images stéréoscopiques "Couloir" provient du site internet ⁵ du groupe *Computer Vision* du département *Computer Science* de l'université de Bonn en Allemagne (cf. figure 4.20). Ces images représentent un couloir, trois objets placés par terre à des distances différentes : une sphère située au milieu et tout près du stéréoscope, un cône placé un peu plus loin que la sphère et à gauche de celle-ci et, tout au bout du couloir, un disque. En dehors de ces objets, on note la présence de quelques tableaux accrochés aux murs.

Les figures 4.20 et 4.21 présentent les images originales gauche et droite ainsi que les résultats de l'extraction des points de contour obtenus sur ce couple d'images stéréoscopiques "Couloir". Sur cette figure on identifie bien les différents objets présents dans le couloir à savoir la sphère, le cône et le disque. Le module a pu également extraire les points de contour représentant les murs ainsi que les tableaux.

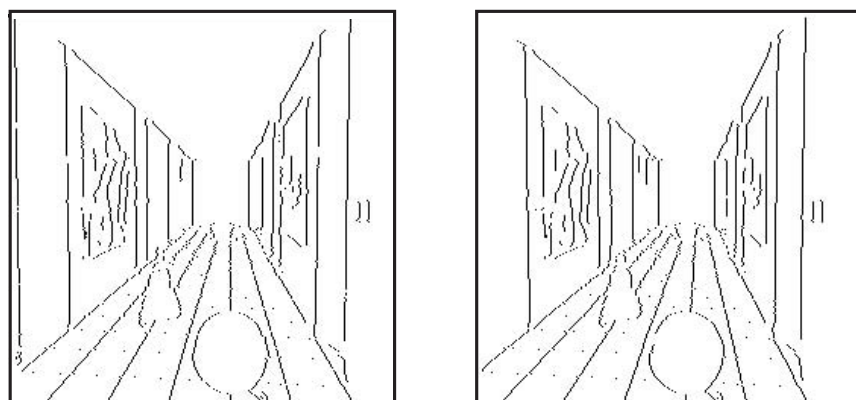
Nous rappelons que l'absence des contours horizontaux sur la figure 4.21 est due à la nature monodimensionnelle du filtre de Deriche appliqué sur chacune des lignes composant l'image matricielle indépendamment des autres.

⁵http://www-dbv.cs.uni-bonn.de/stereo_data



(a) Image gauche.

(b) Image droite.

FIG. 4.20 : Couple d'images "Couloir".

(a) Points de contour extraits de l'image gauche.

(b) Points de contour extraits de l'image droite.

FIG. 4.21 : Extraction de contours du couple d'images "Couloir".

4.10.2.2 Résultat de la mise en correspondance

La figure 4.22 montre la scène reconstruite après avoir appliqué la procédure génétique de mise en correspondance. Comme dans l'image de la reconstruction 3D du couple d'images "Voie", la distance par rapport au stéréoscope pour chaque couple de points de contour appariés est représentée en couleur. Celle-ci varie du rouge, qui correspond à la distance la plus petite par rapport au stéréoscope, au bleu qui correspond à la distance la plus grande.

Sur cette figure, on constate que les points de contour des dalles du couloir sont bien appariés sur l'ensemble des images lignes qui les représentent. En effet, les couleurs correspondant à la distance des pavés varient du rouge au vert foncé, ce qui prouve que leur distance s'accroît

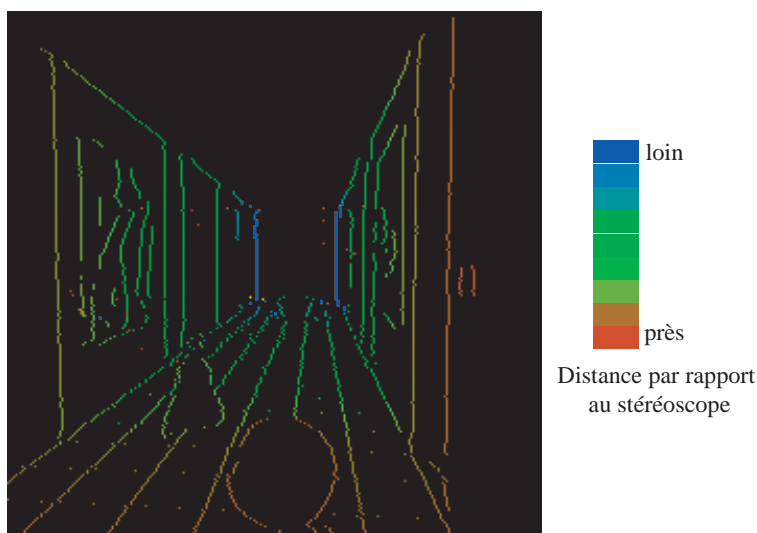


FIG. 4.22 : Résultat de la mise en correspondance du couple d'images "Couloir"

progressivement par rapport au stéréoscope. Les points de contour correspondant à la sphère sont aussi bien appariés. En effet, l'image de la reconstruction montre qu'elle est située tout au début du couloir avec une couleur rouge indiquant sa proximité du stéréoscope. On aperçoit aussi le cône situé à gauche de la sphère avec une couleur verte traduisant le fait qu'il est un peu plus loin que la sphère par rapport au stéréoscope. Le disque situé tout au fond du couloir est aussi bien positionné avec une couleur vert foncé indiquant son éloignement. Outre les objets présents sur le dallage du couloir, la procédure développée a permis d'apparier les points de contour correspondant aussi bien aux murs qu'aux tableaux fixés sur ces derniers. Sur cette figure, on note aussi la présence de quelques mauvais appariements, mais leur nombre est très limité.

Le temps du traitement du couple "Couloir", composé de 256 d'images-ligne stéréoscopiques, est de l'ordre de 80 secondes, c'est-à-dire 0.3 seconde environ par couple d'images-ligne.

4.10.3 Traitement du couple d'images "Village"

4.10.3.1 Description et extraction des points de contour

Le couple d'images stéréoscopiques "Village" a été téléchargé à partir du site internet du laboratoire *CIL (Calibrated Imaging Laboratory)* de l'université *Carnegie Mellon, Pittsburgh, USA*. Ces images sont prises avec une seule caméra qui peut être déplacée grâce à une plateforme motorisée pour obtenir différentes prises de vues. Le lecteur intéressé pourra trouver sur

le site du laboratoire ⁶ plus de détails sur la méthode de calibration, le mouvement de la caméra ainsi que d'autres informations.

Comme pour les autres images de synthèse, les images gauche et droite du couple "Village" sont choisies de sorte que les points images en correspondance diffèrent seulement par leurs coordonnées horizontales, comme s'il s'agissait de caméras placées dans la configuration particulière adoptée dans le cadre de ce travail (cf. figure 2.2).

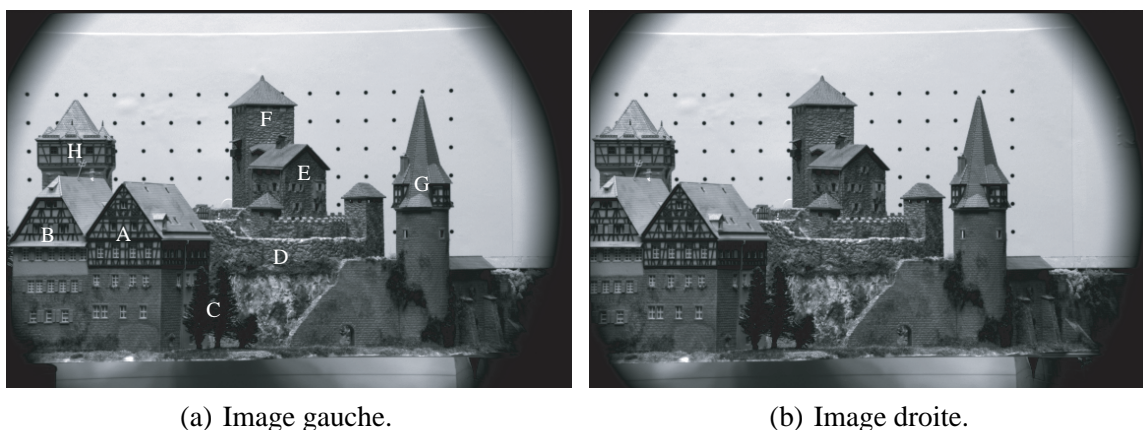


FIG. 4.23 : Couple d'images "Village".

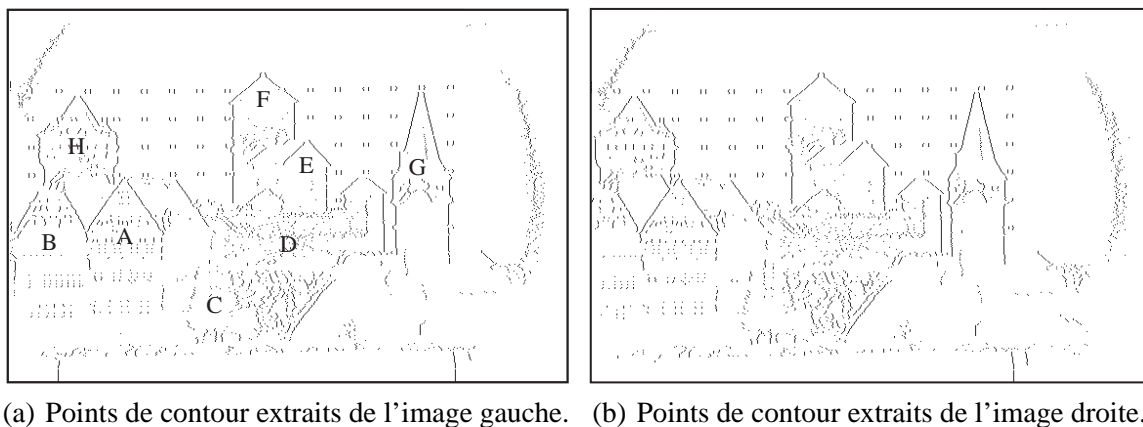


FIG. 4.24 : Extraction de contours du couple d'images "Village".

Les images de la figure 4.23 représentent une scène composée en premier plan de deux maisons (A,B) à gauche de l'image et de quelques arbres (C). Derrière la maison (B), on aperçoit une tour (H), et à sa droite, un château constitué d'un rempart (D), d'une maison (E) et, en arrière plan de la scène, d'un bâtiment (F). À côté du château et à sa droite on trouve une autre tour (G). Sur les images, la grille composée de points noirs est utilisée pour le calibrage de la caméra. De

⁶<http://www-2.cs.cmu.edu/afs/cs/project/cil/ftp/html/cil-ster.html>

plus, on aperçoit des zones sombres au quatre coins des images. Celles-ci sont dues au fait que l'objectif de la caméra focalise la lumière dans une régions circulaire de diamètre inférieur à la surface du capteur CCD de la caméra.

La figure 4.24 montre les résultats de l'extraction des points de contour obtenus sur ce couple d'images stéréoscopiques "Village". Sur cette figure on identifie bien les différents infrastructures décrites plus haut.

4.10.3.2 Résultat de la mise en correspondance

La figure 4.25 montre la reconstruction obtenue grâce à la procédure développée. Afin de permettre une bonne interprétation de l'image de la reconstruction 3D, la distance d'un point de la scène par rapport au stéréoscope est représentée en couleur. Celle-ci varie pour ce couple, du rouge, qui correspond à la distance la plus petite par rapport au stéréoscope, au blanc qui correspond à la distance la plus grande. En effet, pour un souci de clarté, nous souhaitons que les points de contour correspondants à la grille de calibrage et aux zones d'ombres et, qui sont situés au-delà du bâtiment (F), aient une couleur blanche.

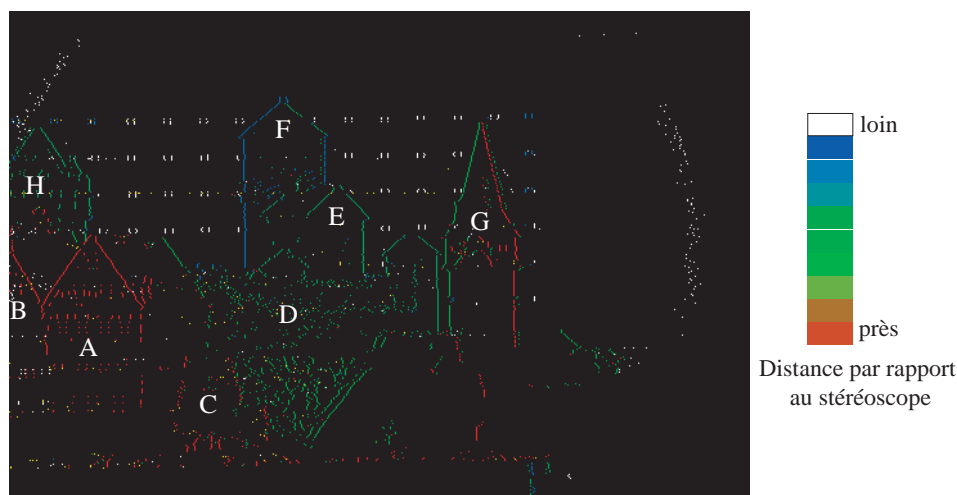


FIG. 4.25 : Résultat de la mise en correspondance du couple d'images "Village"

Le résultat de la reconstruction 3D montre que les points de contour correspondants aux maisons (A) et (B) ainsi qu'aux arbres (C) qui sont situés en premier plan de la scène, sont bien appariés. En effet, les maisons (A) et (B) ainsi que les arbres (C) situés à gauche de la scène ont des couleurs rouge qui sont cohérentes avec leur disposition spatiale car ils sont les plus proches du stéréoscope. En dehors des deux maisons et des arbres, la procédure génétique de mise en correspondance a bien réussi à appairer les points de contour correspondant aux

différentes infrastructures (maison E, tours H et G, rempart D et le bâtiment F) qui font partie du paysage. En effet, les différentes infrastructures ont des couleurs parfaitement cohérentes avec leur disposition spatiale. Sur la figure 4.25, on aperçoit également les points correspondant à la mire qui se trouve tout au fond de la scène. Ces points sont de couleur blanche, ce qui prouve leur éloignement du stéréoscope. Malgré le grand nombre de points de contour candidats à l'appariement extraits de ce couple d'images, la présence de mauvais appariements reste limitée.

Le temps du traitement du couple "Village", composé de 384 d'images-ligne stéréoscopiques, est de l'ordre de 1600 secondes, c'est-à-dire 4 secondes environ par couple d'images-ligne.

Dans le paragraphe suivant, nous traitons un couple d'images stéréoscopiques réelles acquises avec un stéréoscope matriciel installé sur une rame du métro VAL de Lille dans le cadre du projet STATUE. Ce couple d'images est appelé "Rails". Comme les autres images stéréoscopiques, celles-ci sont traitées par le même module d'extraction de primitives. Ensuite les deux listes de points de contour issues de cette étape sont mises en correspondance par l'algorithme génétique décrit précédemment.

4.10.4 Traitement du couple d'images "Rails"

4.10.4.1 Description et extractions des points de contour

Les images stéréoscopiques traitées dans ce paragraphe ont été acquises sur le site d'essais du VAL ⁷ grâce à un capteur stéréoscopique (cf. figure 4.26) installé sur une rame du métro (cf. figure 4.27). Ces images montrent une voie et différents objets faisant office d'obstacles (pigeons, ballon, parpaings, etc.) en présence d'une rame en approche. D'autres structures apparaissent également sur les images telles que le figurant, le feu rouge, les capteur de passages, etc.

Les figures 4.28(a) et 4.28(b) montrent que l'extraction des contours permet de bien identifier les rails du métro ainsi que les différents objets présents sur le site.

4.10.4.2 Résultat de la mise en correspondance

La figure 4.29 montre le résultat de la mise en correspondance obtenu grâce à la procédure génétique développée. Comme auparavant, pour permettre une bonne interprétation du résultat de la mise en correspondance de ce couple d'images, la distance d'un point de la scène réelle

⁷Véhicule Automatique Léger

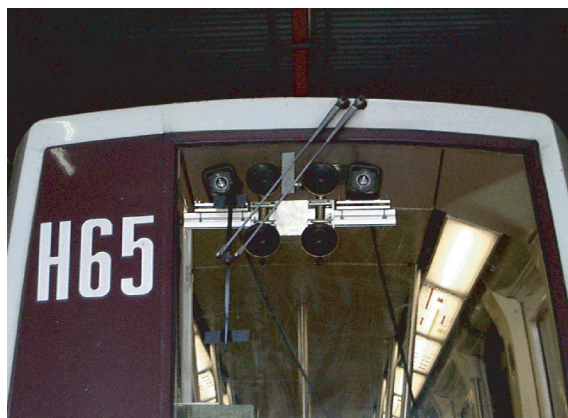


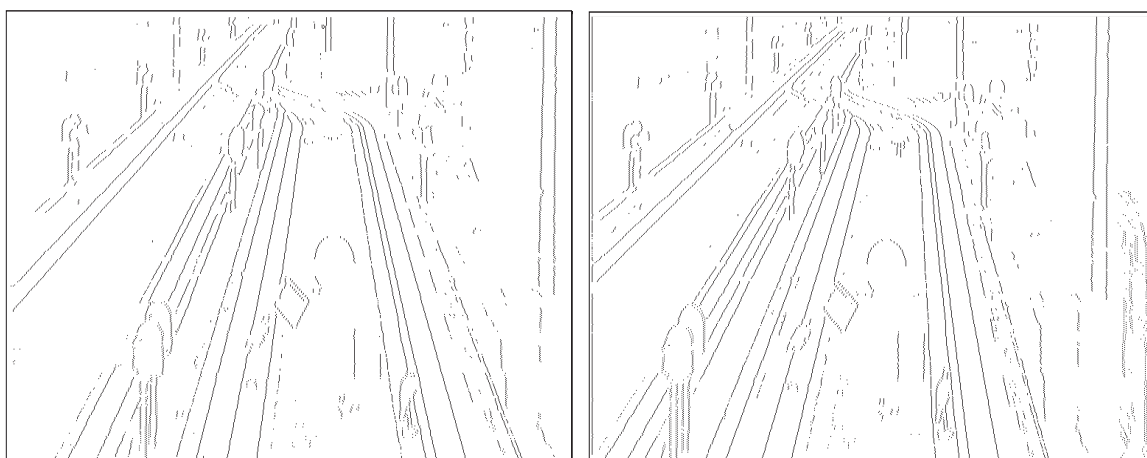
FIG. 4.26 : Stéréoscope dans la rame.



(a) Image gauche.

(b) Image droite.

FIG. 4.27 : Couple d'images "Rails".



(a) Points de contour extraits de l'image gauche.

(b) Points de contour extraits de l'image droite.

FIG. 4.28 : Extraction de contours du couple d'images "Rails".

par rapport au stéréoscope est représentée en couleur. Celle-ci varie du rouge, qui correspond à la distance la plus petite par rapport au stéréoscope, au bleu qui correspond à la distance la plus grande.

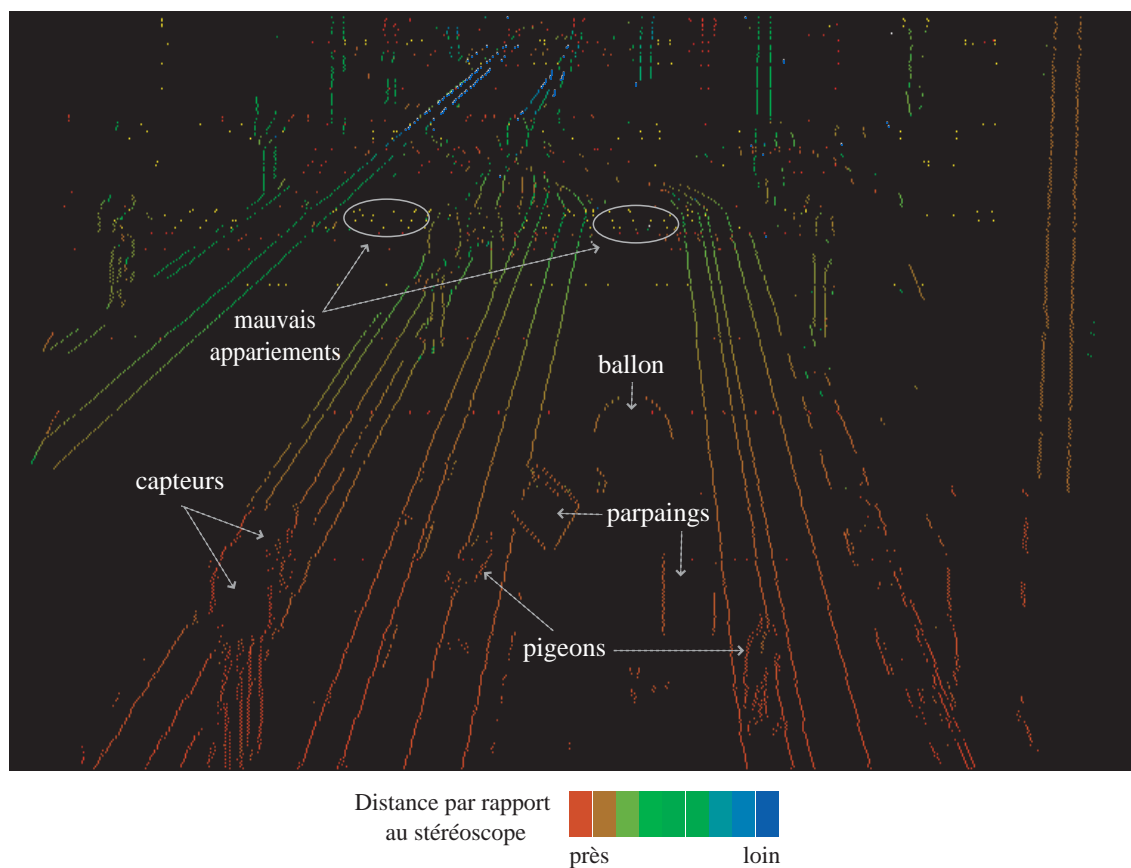


FIG. 4.29 : Résultat de la mise en correspondance du couple d'images "Rails".

Le résultat de la mise en correspondance montre que les points de contour correspondant aux rails du métro sont bien appariés et ceci sur l'ensemble des images-ligne où ils apparaissent. En effet, la couleur à ces points varie progressivement du rouge au blanc tout en passant par le vert et le bleu ce qui montre bien la continuité des rails dans la scène observée. Les points de contour correspondant aux différents objets déposés sur la voie sont aussi bien appariés. Sur l'image de la reconstruction, on identifie facilement les pigeons, les parpaings et le ballon. Les couleurs qui leur sont attribuées indiquent qu'ils sont situés près du stéréoscope (les pigeons sont les plus proches du stéréoscope puis viennent les parpaings et ensuite le ballon). En dehors des rails du métro et des objets déposés sur la voie, la procédure génétique a globalement bien apparié les points de contour correspondant aux différentes structures présentes dans la scène réelle, tels que les capteurs de détection de passage, le figurant, le feu rouge, etc... On remarque

cependant la présence de mauvais appariements que l'algorithme n'a pas pu éviter. Ces mauvais appariements concernent globalement les points de contour appartenant aux structures éloignées du stéréoscope et qui sont situées très près les unes des autres. Nous pensons qu'une étude plus approfondie permettrait de mieux régler les différents paramètres intervenant dans la procédure génétique développée.

Le temps du traitement du couple "Rails", composé de 425 d'images-ligne stéréoscopiques, est de l'ordre de 4400 secondes, c'est-à-dire 10 secondes environ par couple d'images-ligne.

4.11 Conclusion

Dans ce chapitre, nous avons proposé un nouveau codage génétique du problème de la mise en correspondance des points de contour extraites d'images stéréoscopiques. Les nouveaux chromosomes créés grâce à ce codage sont composés de gènes prenant des valeurs entières plutôt que binaires. Ceci a permis, d'une part, de réduire la taille des chromosomes et, d'autre part, d'explorer plus efficacement l'espace de recherche de solutions grâce à l'élimination des gènes représentant des appariements ambigus. Ainsi, la probabilité de converger vers une solution acceptable est devenue plus grande.

En dehors de ces deux avantages, le nouveau codage a permis de réduire d'un facteur 1.5 le temps de traitement de la séquence "Piéton". Cependant, ce temps reste assez important quand l'évaluation de l'adaptation des chromosomes entiers s'effectue à partir des chromosomes binaires correspondants. Il nous a alors paru judicieux de mettre à jour la fonction d'évaluation pour qu'elle permette de mesurer directement l'adaptation des chromosomes entiers sans passer par les chromosomes binaires correspondants. La nouvelle fonction d'évaluation a permis de réduire le temps de traitement d'un facteur de 10. En d'autres termes, le codage entier et la nouvelle fonction d'évaluation ont permis, d'une part, d'explorer plus efficacement l'espace de solutions et, d'autre part, de réduire le temps de traitement d'un facteur de 15 environ par rapport au premier codage présenté dans le chapitre précédent.

Ensuite, le codage entier et la nouvelle fonction d'évaluation ont été utilisés pour mettre en correspondance différents couples d'images stéréoscopiques de synthèse et réelles. Ces images sont créées à partir d'un stéréoscope placé dans la configuration particulière permettant de confondre les droites épipolaires avec les lignes des capteurs. Grâce à cette configuration particulière, nous avons pu mettre en correspondance ces images sans faire appel à des procédures

supplémentaires. Les résultats de la mise en correspondance obtenus lors des essais effectués sur les couples d'images de synthèse et sur le couple d'images réelles sont satisfaisants mais le temps de traitement nécessaire pour l'appariement de certains couples d'images-ligne stéréoscopiques reste trop important, notamment pour des applications nécessitant de réponses en temps réel tel que la détection d'obstacles. En effet, lorsque les images-ligne stéréoscopiques contenant un nombre important de points de contour candidats à l'appariement, la taille des chromosomes entiers augmente et la convergence vers une solution acceptable nécessite l'augmentation de la taille des populations et du nombre de générations. De ce fait, notre approche globale de mise en correspondance nécessite un temps de calcul afin d'apparier ces images-ligne avant de fournir de résultats satisfaisants, et ceci malgré le gain du temps réalisé grâce au codage entier.

Dans le chapitre suivant, nous décrivons une nouvelle méthode de mise en correspondance d'images-ligne stéréoscopiques combinant deux approches différentes. La première est l'approche génétique développée précédemment et la deuxième est une approche hiérarchique permettant d'effectuer une recherche des appariements à différents niveaux. L'objectif de cette nouvelle méthode est de réduire la taille des chromosomes entiers afin de diminuer encore plus le coût des calculs nécessaires à leur évaluation.

Chapitre 5

Stratégie de recherche hiérarchique pour la mise en correspondance stéréoscopique

5.1 Introduction

Nous avons montré dans le chapitre précédent que le codage entier est plus efficace que le codage binaire dans le sens où le premier permet à un algorithme génétique de mieux explorer l'espace des solutions. En effet, avec l'utilisation du codage entier, l'algorithme génétique fournit des solutions meilleures tant au niveau qualité qu'au niveau coût des calculs en termes de nombre de générations et de taille de la population. Par ailleurs, le codage entier fournit des chromosomes de taille moins importante, ce qui permet, grâce à une fonction d'évaluation spécifique, de réduire le coût des calculs nécessaires à leur évaluation. Malgré ces améliorations considérables du coût des calculs de notre algorithme génétique, notre procédure de mise en correspondance reste toujours incompatible avec une application temps réel telle que la détection d'obstacles à l'avant d'un véhicule routier. Cette limitation est due principalement à la combinatoire élevée du problème de mise en correspondance. Afin d'envisager une exploitation effective de notre procédure de mise en correspondance pour la détection d'obstacles, nous proposons dans ce chapitre une méthode hiérarchique permettant de décomposer le problème en plusieurs sous problèmes ayant des combinatoires moins élevées. L'idée principale de cette méthode est d'effectuer une recherche des appariements par niveau. Pour un niveau de recherche donné, les points de contour les plus significatifs sont d'abord appariés pour obtenir des couples de référence. Ces derniers vont ensuite servir de base pour mettre en correspondance les points de contour moins significatifs.

5.2 Réduire la combinatoire du problème

Le coût des calculs dans un algorithme génétique dépend généralement de trois paramètres : la taille de la population, le nombre de générations et la taille des chromosomes. Les deux premiers paramètres influent sur la qualité de la solution issue du processus d'évolution induit par l'algorithme génétique. En effet, plus ces deux paramètres prennent des valeurs élevées plus la probabilité d'arriver à une bonne solution augmente, mais avec une augmentation considérable du coût des calculs. Ainsi, agir directement sur ces deux paramètres pour réduire le coût des calculs pourrait conduire à des solutions moins bonnes, ce qui est parfois acceptable dans certains problèmes, mais pas dans le cas où il s'agit de mettre en correspondance des images stéréoscopiques pour détecter des obstacles à l'avant d'un véhicule routier. Le troisième paramètre est lié à la fois au codage et à la combinatoire du problème. Nous avons vu au chapitre précédent que la taille des chromosomes issus du codage entier est moins importante que la taille des chromosomes fournis par le codage binaire. Quand à la relation entre la taille des chromosomes et la combinatoire du problème, il est évident que plus la combinatoire du problème est élevée plus la taille des chromosomes est importante. L'expérience montre que la taille de la population et le nombre de générations sont liés à la taille des chromosomes. En effet, lorsque la taille des chromosomes est moins importante, des valeurs moins élevées de la taille de la population et du nombre de générations suffisent pour que l'algorithme génétique aboutisse à de bonnes solutions. Ainsi, en réduisant la combinatoire du problème, donc la taille des chromosomes, le coût des calculs en termes de nombre de générations et de taille de la population sera réduit. Cependant, cette réduction de la combinatoire nécessite une décomposition du problème en sous-problèmes moins complexes. Cette décomposition doit être effectuée avec précaution pour ne pas trop perdre, lors de la résolution des sous-problèmes, l'apport des contraintes globales qui sont à la base de la méthode de mise en correspondance des images. Afin de réduire la combinatoire du problème de mise en correspondance, nous proposons une méthode de décomposition basée sur une stratégie de recherche hiérarchique. L'idée est de constituer, dans un premier temps, des couples de référence à partir de la mise en correspondance d'un certain nombre de points de contour sélectionnés selon un critère dans les deux images. Ces couples de références servent ensuite de base pour mettre en correspondance les autres points de contours. La suite de ce chapitre est consacrée à la description détaillée de notre méthode hiérarchique de mise en correspondance.

5.3 Analyse du problème

Pour analyser le problème et expliquer notre méthode de recherche hiérarchique, nous faisons appel à la notion de schéma de mise en correspondance que nous avons décrite dans le § 3.3. Considérons le schéma de mise en correspondance de la figure 5.1. L'analyse de cet exemple conduit à la constatation suivante. Pour réduire la complexité du problème, au lieu de mettre en correspondance d'un seul coup tous les points de contour des deux images, il est plus simple de commencer par mettre en correspondance les points de contour les plus significatifs dans les deux images afin d'obtenir des couples de référence à partir desquels les autres points de contour seront ensuite mis en correspondance. Les couples de référence vont permettre alors de générer de nouveaux schémas de mise en correspondance, de tailles inférieures, auxquelles on appliquera le même principe. De cette manière, on décompose le problème de mise en correspondance en plusieurs sous-problèmes ayant des combinatoires moins élevées.

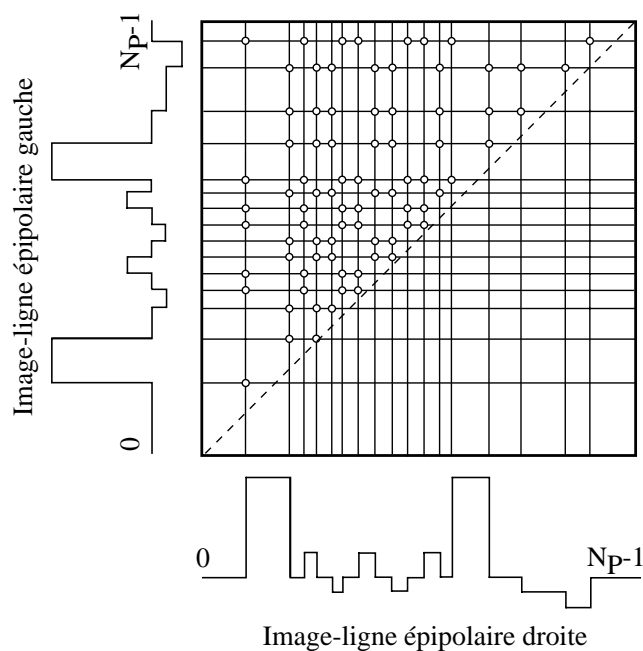


FIG. 5.1 : Schéma de mise en correspondance du problème initial.

L'idée qui ressort de l'analyse de cet exemple est de mettre en correspondance les points de contour hiérarchiquement grâce à l'utilisation de couples de référence. L'obtention de ces points de référence nécessite la recherche, dans chaque image, de points de contour plus significatifs que les autres. Or, la notion de points de contour significatifs sous-entend les classements, dans chaque image, des points de contour en fonction d'un critère de telle sorte que ces classements soient compatibles entre eux, c'est-à-dire que les points de contour homologues soient classés

dans le même ordre. Cette compatibilité est très importante car, sans elle, l'idée de la recherche hiérarchique devient inutilisable du fait que les risques d'erreurs d'appariement vont augmenter. En effet, avec un mauvais critère de sélection, on peut se trouver dans des situations où un point de contour est sélectionné dans une image alors que son correspondant dans l'autre image ne l'est pas. Dans ce cas, lors de la mise en correspondance des points sélectionnés, des faux appariements vont se produire mais également vont se propager puisque la mise en correspondances des autres points de contour en dépend.

5.4 Principe de base de la méthode hiérarchique

Pour expliquer notre méthode hiérarchique de mise en correspondance, nous utilisons la notion de niveau de recherche. Pour un niveau de recherche donné, on sélectionne dans chaque image un certain nombre de points de contours. La figure 5.2 montre le schéma de mise en correspondance correspondant au niveau 1 de recherche où 4 points de contour ont été sélectionnés dans chaque image. Le premier sous-problème issu de la décomposition consiste alors à mettre en correspondance ces points de contours sélectionnés. On constate que la combinatoire du premier sous-problème est moins importante. Nous appliquons ensuite notre algorithme génétique de mise en correspondance pour appairer les points de contours sélectionnés.

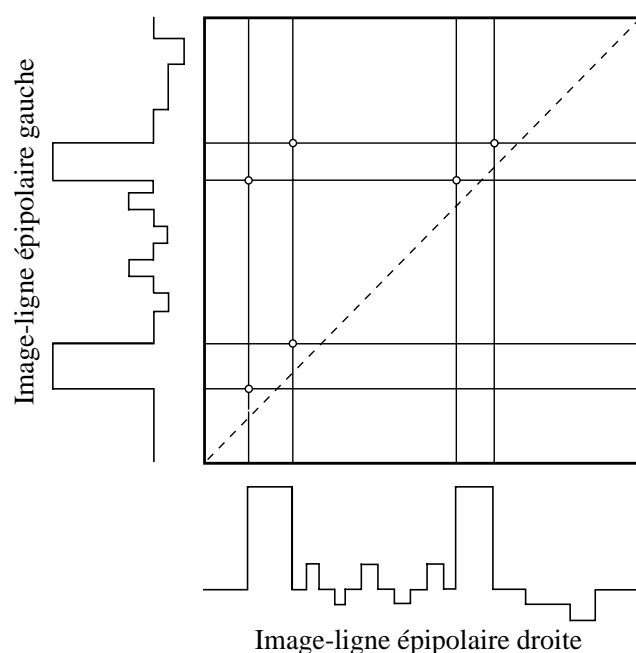


FIG. 5.2 : Niveau 1 de recherche.

A l'issue de la procédure de mise en correspondance, nous obtenons une liste de couples

de points de contour appariés. Cette liste sert de base pour mettre en correspondance les autres points de contour au niveau suivant. La figure 5.3 montre le niveau 2 de recherche. Deux couples successifs servent à définir un nouveau schéma de mise en correspondance. Les nouveaux schémas situés aux extrémités du schéma initial sont aussi considérés au niveau 2 de recherche. Dans notre exemple, nous pouvons voir 5 nouveaux schémas de mise en correspondance définies à partir des 4 couples de référence. Comme au premier niveau de recherche, les schémas du second niveau de recherche représentent des sous-problèmes de mise en correspondance avec des complexités moins importantes.

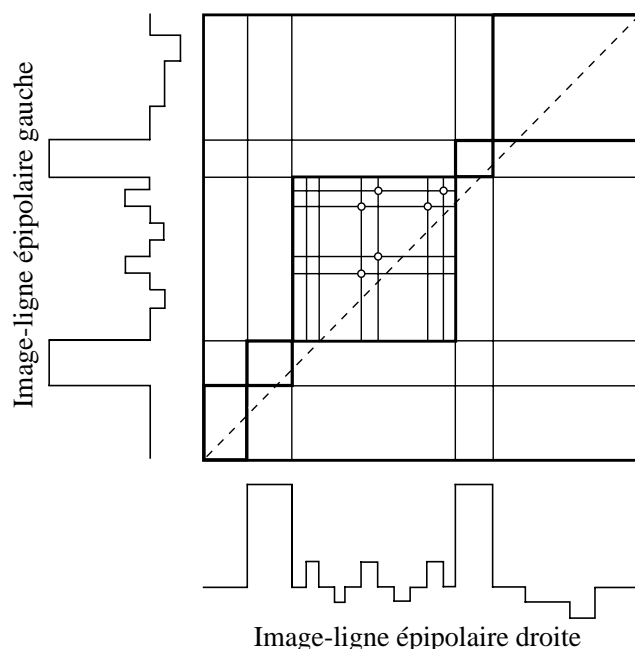


FIG. 5.3 : Niveau 2 de recherche.

Au niveau 2 de recherche, on applique le même procédé en considérant chacun des nouveaux schémas de mise en correspondance. Pour chaque schéma, on sélectionne, parmi les points de contour qui lui sont associés, les points de contour les plus significatifs. Dans le cas de la figure 5.3, un seul schéma de mise en correspondance sera traité au niveau de 2 de recherche. On applique alors l'algorithme génétique de mise en correspondance pour appairer les points de contour sélectionnés. Les couples appariés obtenus servent ensuite de base pour définir de nouveaux schémas représentant des sous-problèmes de mise en correspondance de combinatoires encore moins importantes. Cette procédure de recherche hiérarchique est implantée de façon récursive.

5.5 Description de la méthode hiérarchique

Avant de se lancer dans la description de la méthode hiérarchique, nous commençons par donner quelques définitions et notations utiles à sa mise en œuvre.

5.5.1 Définitions

Un schéma de mise en correspondance est construit à partir d'une liste de points de contour de l'image-ligne gauche et une liste de points de contour de l'image-ligne droite. Une liste de points de contour d'une image est constituée de l'ensemble des points de contour de l'image-ligne qui sont situés entre deux points de l'image. Ainsi, un schéma de mise en correspondance est obtenu à partir de deux couples de points. Chaque couple, que nous appelons couple de référence, est constitué d'un point de l'image-ligne gauche et d'un point de l'image-ligne droite. Pour pouvoir commencer le traitement, on définit le schéma de mise en correspondance au niveau 1 de recherche grâce aux deux couples de référence $(0, 0)$ et $(N_P - 1, N_P - 1)$ où N_P est le nombre de pixels dans chaque image-ligne. Ces deux couples de référence seront 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 créer un schéma de mise en correspondance au niveau supérieur. Notons $CR_n^{q'}(q) = (g_n^{q'}(q), d_n^{q'}(q))$ le $q^{\text{ème}}$ couple de référence obtenu à partir du schéma de rang q' au niveau n où $g_n^{q'}(q)$ et $d_n^{q'}(q)$ sont respectivement deux points des images gauche et droite. Ce couple servira donc à la création de nouveaux schémas de mise en correspondance au niveau $n + 1$.

Soit

$$M_n^{q'}(q) = [CR_{n-1}^{q'}(q-1), CR_{n-1}^{q'}(q)]$$

le schéma de mise en correspondance de rang q au niveau n de recherche. Il est défini par les deux couples de référence successifs

$$CR_{n-1}^{q'}(q-1) = (g_{n-1}^{q'}(q-1), d_{n-1}^{q'}(q-1))$$

et

$$CR_{n-1}^{q'}(q) = (g_{n-1}^{q'}(q), d_{n-1}^{q'}(q))$$

obtenus à partir du schéma de rang q' au niveau $n - 1$. Par conséquent, $M_n^{q'}(q)$ représente le

schéma de mise en correspondance de rang q au niveau n , obtenu à partir du schéma de rang q' au niveau $n - 1$.

Soient $G_n^{q'}(q)$ et $D_n^{q'}(q)$ les listes de points de contour gauches et droits permettant de construire le schéma $M_n^{q'}(q)$. La liste $G_n^{q'}(q)$ est constituée de l'ensemble des points de contour de l'image-ligne gauche situés entre les points $g_{n-1}^{q'}(q-1)$ et $g_{n-1}^{q'}(q)$. De même, la liste $D_n^{q'}(q)$ est constituée de l'ensemble des points de contour situés entre les points $d_{n-1}^{q'}(q-1)$ et $d_{n-1}^{q'}(q)$. Le traitement du schéma $M_n^{q'}(q)$ consiste, dans un premier temps, à sélectionner, à partir des listes $G_n^{q'}(q)$ et $D_n^{q'}(q)$, les points de contour les plus significatifs. Notons $GS_n^{q'}(q)$ et $DS_n^{q'}(q)$ les listes des points de contour sélectionnés à partir des listes $G_n^{q'}(q)$ et $D_n^{q'}(q)$, respectivement. Les listes $GS_n^{q'}(q)$ et $DS_n^{q'}(q)$ sont ensuite mises en correspondance pour constituer de nouveaux couples de référence à partir desquels de nouveaux schémas seront générés au niveau $n + 1$. Notons $CR_n^q = \{CR_n^q(k), k = 0, \dots, K_n^q\}$ la liste des $K_n^q + 1$ couples de références obtenus à l'issue du traitement du schéma $M_n^{q'}(q)$. En plus des couples de points de contour appariés fournis par la procédure de mise en correspondance des listes $GS_n^{q'}(q)$ et $DS_n^{q'}(q)$, cette liste contient également les couples de référence $CR_{n-1}^{q'}(q-1)$ et $CR_{n-1}^{q'}(q-1)$ qui ont été utilisés pour définir le schéma $M_n^{q'}(q)$. Ces deux couples sont nécessaires à la création des nouveaux schémas de mise en correspondance au niveau $n + 1$, notamment le premier et le dernier schémas. Précisons que les couples de référence de la liste $CR_n^q(q)$ sont classés selon l'ordre des points de contour dans les images-ligne. Ainsi, le premier couple $CR_n^q(0)$ (resp. dernier couple $CR_n^q(K_n^q)$) de la liste $CR_n^q(q)$ n'est autre que le couple $CR_{n-1}^{q'}(q-1)$ (resp. $CR_{n-1}^{q'}(q)$).

A partir de la liste CR_n^q des couples de référence, on génère alors au niveau $n + 1$ les K_n^q schémas de mise en correspondance obtenus à partir du schéma de rang q au niveau n .

Notons $M_{n+1}^q = \{M_{n+1}^q(k), k = 0, \dots, K_n^q\}$ la liste de ces schémas où

$$M_{n+1}^q(k) = [CR_n^q(k-1), CR_n^q(k)]$$

est le schéma de mise en correspondance de rang k au niveau $n + 1$, obtenu à partir des $(k-1)^{\text{ème}}$ et $k^{\text{ème}}$ couples de référence issus du traitement du schéma de rang q au niveau n .

5.5.2 Algorithme génétique de mise en correspondance et méthode hiérarchique

L'algorithme génétique de mise en correspondance que nous avons présenté dans le chapitre 3 s'applique sur l'ensemble des points de contour extraits des images gauche et droite.

Un seul schéma est utilisé pour représenter la mise en correspondance des deux images-ligne épipolaires. Il est construit à partir de l'ensemble des points de contour des images-ligne épipolaires gauche et droite. En tenant compte des définitions données ci-dessus, ce schéma peut être obtenu à partir des deux couples de référence $CR_0^1(0) = (0, 0)$ et $CR_0^1(1) = (N_p - 1, N_p - 1)$. Il est alors noté par $M_1^1(1) = [CR_0^1(0), CR_0^1(1)]$. Par conséquent, la notation $M_1^1(1)$ représente le schéma de mise en correspondance de rang 1 au niveau 1, obtenu à partir du schéma fictif de rang 1 au niveau 0.

Ainsi, l'algorithme génétique de mise en correspondance peut être vu comme un algorithme hiérarchique particulier avec un seul niveau de recherche où tous les points de contour sont pris en compte. Celui-ci pourra donc être paramétré afin de considérer tout schéma de mise en correspondance définie par deux couples de référence.

5.6 Implantation récursive de la méthode hiérarchique

5.6.1 Principe de base

La méthode hiérarchique est implantée selon un algorithme récursif afin de traiter tous les schémas de mise en correspondance, quelque soit leur nombre. Pour expliquer cette procédure récursive, 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 d'un élément d'une paire vers l'autre. 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 problème, les nœuds de l'arbre sont les schémas de mise en correspondance. Le schéma de mise en correspondance du niveau 1 de recherche constitue la racine de l'arbre. A partir d'un nœud du niveau n , c'est-à-dire un schéma du niveau n , l'application du module d'appariement élémentaire hiérarchique (cf. § 5.6.4.4) génère des nœuds fils au niveau $n + 1$.

Un exemple d'arbre expliquant le fonctionnement de l'algorithme récursif est représenté sur la figure 5.4. Dans cet exemple, le nœud $M_1^1(1) = [CR_0^1(0), CR_0^1(1)]$ du niveau 1 de recherche est défini à partir des deux couples de référence $CR_0^1(0) = (0, 0)$ et $CR_0^1(1) = (N_p - 1, N_p - 1)$. L'application du module d'appariement élémentaire hiérarchique au nœud $M_1^1(1)$ a permis d'apparier 4 couples de points de contour. Ainsi, les couples de référence $CR_1^1(0) = CR_0^1(0)$, $CR_1^1(1)$, $CR_1^1(2)$, $CR_1^1(3)$, $CR_1^1(4)$, $CR_1^1(5) = CR_0^1(1)$ permettent de

créer les nouveaux nœuds $M_2^1(1) = [CR_1^1(0), CR_1^1(1)]$, $M_2^1(2) = [CR_1^1(1), CR_1^1(2)]$, $M_2^1(3) = [CR_1^1(2), CR_1^1(3)]$, $M_2^1(4) = [CR_1^1(3), CR_1^1(4)]$, $M_2^1(5) = [CR_1^1(4), CR_1^1(5)]$.

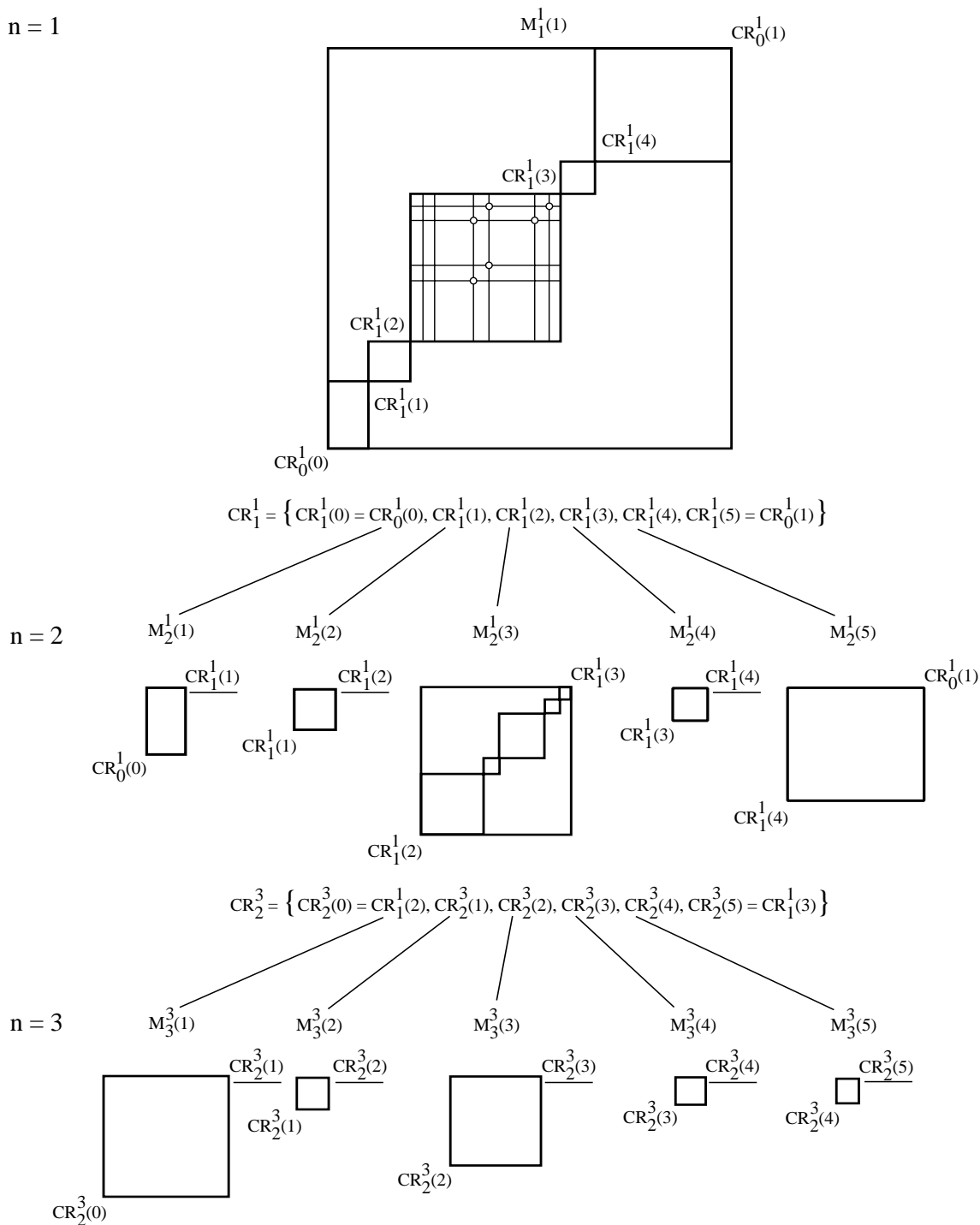


FIG. 5.4 : Arbre de recherche de la stratégie hiérarchique.

5.6.2 Exploration de l'arbre

Pour trouver l'ensemble des appariements, la procédure consiste à développer les nœuds de l'arbre en utilisant une stratégie d'exploration dite en profondeur d'abord, qui consiste à traiter les nœuds de niveaux supérieurs avant de traiter les nœuds de même niveau. Lorsque tous les nœuds d'un même niveau ont été 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 cherche d'abord à étendre les recherches au niveau $n + 2$. Pour cela, une priorité est toujours donnée au nœud fils le plus à gauche. Un test d'arrêt à deux conditions est effectué pour déterminer le type de ce nœud (cf. § 5.6.4.3). En fonction du résultat du test, le nœud peut être de type terminal, post-terminal ou normal (cf. § 5.6.4.3). Si le nœud est de type terminal, il cesse d'engendrer de nouveaux nœuds au niveau $n + 2$. 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 hiérarchique sans sélection. Celui-ci effectue la procédure de mise en correspondance sur l'ensemble des points de contour relatifs au nœud. Si ce module fournit de nouvelles paires appariées, celles-ci généreront au niveau $n + 2$ de nouveaux nœuds fils qui sont des nœuds feuille. Sinon, le nœud lui-même devient un nœud feuille. Dans le cas où le nœud est de type normal, on lance le module d'appariement élémentaire hiérarchique avec sélection. Celui-ci consiste à appliquer la procédure de mise en correspondance aux points de contour les plus significatifs du nœud. Si ce module fournit de nouvelles paires appariées, celles-ci généreront au niveau $n + 2$ de nouveaux nœuds fils à partir desquels le processus est itéré en commençant par le nœud le plus à gauche. Sinon, on re-traite le même nœuds au niveau $n + 1$. Lorsque le module d'appariement élémentaire hiérarchique avec sélection ne fournit pas de nouvelles paires, cela signifie que l'une des listes des points de contour sélectionnés est vide ou tous simplement que la procédure de mise en correspondance n'a pas réussi à trouver d'appariement. Dans les deux cas, nous avons décidé de continuer la recherche dans le même nœud au niveau supérieur pour donner une chance à la fois à la procédure de sélection (cf. § 5.6.4.1) afin de fournir des listes de points de contour non vides et à la procédure de mise en correspondance pour trouver des appariements.

Lorsque le processus cesse d'engendrer de nouveaux nœuds dans l'arbre, les nœuds feuille fournissent alors les résultats de mise en correspondance. Pour cela, il suffit de considérer les

deux couples de référence définissant chaque nœud feuille et de conserver la paire correspondant au deuxième couple (cf. § 5.6.3).

Pour comprendre ce procédé, il suffit de regarder l'arbre de la figure 5.4 qui traite l'exemple de la figure 5.1. Dans cet arbre, le nœud racine $M_1^1(1)$ génère successivement les nœuds $M_2^1(1)$, $M_2^1(2)$, $M_2^1(3)$, $M_3^3(1)$, $M_3^3(2)$, $M_3^3(3)$, $M_3^3(4)$, $M_3^3(5)$, $M_2^1(4)$, $M_2^1(5)$. On commence par appliquer le module d'appariement élémentaire hiérarchique avec sélection au nœud $M_1^1(1)$ défini par les deux couples de référence $CR_0^1(0)$ et $CR_0^1(1)$. A l'issue du traitement de ce nœud, 4 paires sont appariées. On constitue alors la liste des couples de référence $CR_1^1 = \{CR_1^1(0) = CR_0^1(0), CR_1^1(1), CR_1^1(2), CR_1^1(3), CR_1^1(4), CR_1^1(5) = CR_0^1(1)\}$ à partir de laquelle on crée au niveau 2 de recherche 5 nouveaux nœuds fils $M_2^1(1)$, $M_2^1(2)$, $M_2^1(3)$, $M_2^1(4)$, $M_2^1(5)$. Comme l'arbre est exploré en utilisant une stratégie en profondeur d'abord, on traite 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 hiérarchique sans sélection. A partir des 4 paires appariées, on constitue la liste $CR_2^3 = \{CR_2^3(0) = CR_1^1(2), CR_2^3(1), CR_2^3(2), CR_2^3(3), CR_2^3(4), CR_2^3(5) = CR_1^1(3)\}$ des couples de référence, ce qui permet alors de générer au niveau 3 de recherche 5 nouveaux nœuds fils $M_3^3(1)$, $M_3^3(2)$, $M_3^3(3)$, $M_3^3(4)$, $M_3^3(5)$. Ces derniers sont automatiquement des nœuds feuille. 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.

5.6.3 Prise en compte des résultats de la procédure récursive

La décomposition du problème de mise en correspondance selon la stratégie hiérarchique et son implantation font que la solution ne se trouve pas sur une feuille de l'arbre. En fait, chaque feuille possède une partie de la solution. Il faudra donc atteindre toutes les feuilles pour déterminer l'ensemble des appariements.

Il existe une autre manière pour déterminer l'ensemble des appariements. Elle consiste à mémoriser les résultats de mise en correspondance issus de chaque appel du module d'appariement élémentaire hiérarchique. Dans le cas de l'exemple de la figure 5.4, nous sauvegardons les résultats du nœud $M_1^1(1)$, puis ceux du nœud $M_2^1(1)$, etc. Dans la liste finale ainsi obtenue, les paires de points de contour sont rangées dans l'ordre où elles sont appariées. Cependant, pour

faciliter l'utilisation ultérieure de cette liste, nous préférons conserver l'ordre que présentent les points de contour dans les images-ligne épipolaires.

Un appariement sera donc retenu quand un nœud ne pourra plus générer de nouveaux nœuds. Pour cela, il suffira de considérer les deux couples de référence définissant chaque feuille de l'arbre et de ne conserver que la paire correspondant au deuxième couple. N'oublions pas toutefois qu'au niveau 1, le schéma de mise en correspondance est initialisé avec les couples de référence $CR_0^1(0) = (0, 0)$ et $CR_0^1(1) = (N_P - 1, N_P - 1)$ constitués par les premiers et les derniers pixels de chaque image. Pour cela, il suffit simplement d'ignorer la dernière feuille de l'arbre.

Dans le cas de l'exemple de la figure 5.4, après avoir traité le nœud $M_1^1(1)$, l'algorithme aboutira au niveau 2 sur les feuilles $M_2^1(1)$ puis $M_2^1(2)$. Il mémorisera alors successivement les paires correspondant aux couples $CR_1^1(1)$ et $CR_1^1(2)$ qui auraient été obtenus à partir du schéma de rang 1 au niveau 1. Au niveau 3, il atteindra ensuite les feuilles $M_3^3(1)$, $M_3^3(2)$, $M_3^3(3)$, $M_3^3(4)$, $M_3^3(5)$ et mémorisera successivement les paires correspondant aux couples $CR_2^3(1)$, $CR_2^3(2)$, $CR_2^3(3)$, $CR_2^3(4)$, $CR_2^3(5)$ qui auraient été obtenus à partir du schéma $M_2^1(3)$ de rang 3 au niveau 2. Finalement l'algorithme sera amené à atteindre la feuille $M_2^1(4)$ et à mémoriser le couple $CR_1^1(4)$ qui aura été obtenu au niveau 1. Sur la figure 5.4, les paires retenues correspondent aux couples soulignés. On constate que la paire $CR_1^1(5)$ du nœud $M_2^1(5)$ n'est pas retenue puisqu'elle correspond à la dernière feuille de l'arbre.

5.6.4 Implantation récursive

Pour pouvoir finaliser l'implantation récursive de notre méthode hiérarchique de mise en correspondance, nous devons développer :

- Le module d'appariement élémentaire hiérarchique.
- La procédure de sélection des points de contour les plus significatifs.

et définir :

- L'initialisation de l'algorithme.
- Les conditions d'arrêt de l'algorithme.

5.6.4.1 Sélection des points de contour les plus significatifs

Soit $M_n^{q'}(q) = [CR_{n-1}^{q'}(q-1), CR_{n-1}^{q'}(q)]$ le schéma de mise en correspondance de rang q au niveau n , obtenu à partir du schéma de rang q' au niveau $n-1$. Rappelons que les couples

de référence $CR_{n-1}^{q'}(q-1) = (g_{n-1}^{q'}(q-1), d_{n-1}^{q'}(q-1))$ et $CR_{n-1}^{q'}(q) = (g_{n-1}^{q'}(q), d_{n-1}^{q'}(q))$ sont obtenus à partir du schéma de rang q' au niveau $n-1$. Lorsque ce schéma ne vérifie pas l'une des deux conditions d'arrêt de l'algorithme (cf. § 5.6.4.3), on commence par sélectionner des points de contour dans chaque image sur lesquels la procédure de mise en correspondance sera appliquée.

L'idée retenue pour effectuer cette sélection est que les contours qui délimitent des régions avec un fort contraste sont plus faciles à mettre en correspondance que les contours délimitant des régions présentant un faible contraste. Nous avons donc décidé de sélectionner les points de contour en fonction de l'amplitude de leur dérivée.

Comme on s'intéresse au schéma de mise en correspondance $M_n^{q'}(q)$, seuls les points de contour des listes $G_n^{q'}(q)$ et $D_n^{q'}(q)$, qui permettent de la construire, sont concernés par la sélection. Rappelons que $G_n^{q'}(q)$ est la liste de l'ensemble des points de contour de l'image-ligne épipolaire gauche situés entre les points $g_{n-1}^{q'}(q-1)$ et $g_{n-1}^{q'}(q)$. De même, $D_n^{q'}(q)$ est la liste de l'ensemble des points de contour de l'image-ligne épipolaire droite situés entre les points $d_{n-1}^{q'}(q-1)$ et $d_{n-1}^{q'}(q)$. Ainsi, les points de contour candidats à la sélection dans l'image-ligne épipolaire gauche sont ceux qui vérifient la relation :

$$x_{g_{n-1}^{q'}(q-1)} < x < x_{g_{n-1}^{q'}(q)} \quad (5.1)$$

où x désigne l'abscisse des points de contour dans l'image-ligne épipolaire. Pour l'image-ligne épipolaire droite, les points de contours candidats à la sélection doivent vérifier la relation :

$$x_{d_{n-1}^{q'}(q-1)} < x < x_{d_{n-1}^{q'}(q)}. \quad (5.2)$$

Appliqué aux listes $G_n^{q'}(q)$ et $D_n^{q'}(q)$, le critère de sélection est basé sur l'amplitude de la dérivée de l'image-ligne épipolaire. Soit $d(x)$ l'amplitude de la dérivée associée au point de contour d'abscisse x . La valeur $d(x)$ étant non signée, nous prenons également en compte le signe de la dérivée en ce point. La procédure de sélection ne conserve donc que les points de contour gauche présentant une dérivée positive vérifiant la relation :

$$\frac{Max_g}{2n} \leq d(x) \leq Max_g \quad (5.3)$$

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

$$Min_g \leq d(x) \leq \frac{Min_g}{2n} \quad (5.4)$$

où Max_g (resp. Min_g) est l'amplitude maximale (resp. minimale) de la dérivée de l'image-ligne épipolaire gauche.

Ainsi, au premier niveau de recherche, seuls les points de contour dont l'amplitude positive de la dérivée est comprise entre $Max_g/2$ et Max_g sont conservés. On remarque qu'au niveau 2 de la recherche, la sélection est effectuée entre $Max_g/4$ et Max_g . Nous donnons ainsi à un point de contour, qui n'a pas trouvé de correspondant lors de la mise en correspondance au niveau n , la possibilité d'être apparié au niveau $n + 1$. Le raisonnement est identique pour les points de contour dont l'amplitude est négative mais cette fois-ci, les points de contour conservés sont ceux dont l'amplitude de la dérivée est comprise entre Min_g et $Min_g/2$.

Pour la sélection des points de contour droits, des relations similaires sont utilisées :

$$\frac{Max_d}{2n} \leq d(x) \leq Max_d, \quad (5.5)$$

$$Min_d \leq d(x) \leq \frac{Min_d}{2n}, \quad (5.6)$$

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

Ainsi, au niveau n de recherche, la procédure de sélection fournit les listes $GS_n^{g'}(q)$ et $DS_n^{g'}(q)$ qui sont respectivement la liste des points de contour gauches vérifiant simultanément les relations données par les équations 5.1 et 5.3 ou 5.1 et 5.4 et la liste des points de contour droits vérifiant simultanément les relations données par les équations 5.2 et 5.5 ou 5.2 et 5.6.

5.6.4.2 Initialisation de l'algorithme

Comme toute implantation récursive, notre algorithme hiérarchique nécessite une phase d'initialisation pour commencer le traitement. Pour cela, il nous faut, pour le niveau 1 de recherche, un schéma initial de mise en correspondance tenant compte de tous les points de contour des deux images-ligne épipolaires. Ce schéma est défini par $M_1^1(1) = [CR_0^1(0), CR_0^1(1)]$ où $CR_0^1(0) = (0, 0)$ et $CR_0^1(1) = (N_P - 1, N_P - 1)$ sont les couples de référence constitués avec le premier et le dernier pixel de chaque image-ligne épipolaire. Nous rappelons que la notation $M_1^1(1) = [CR_0^1(0), CR_0^1(1)]$ désigne le schéma de mise en correspondance de rang 1 au niveau 1, obtenu à partir du schéma fictif de rang 1 au niveau 0. Notons également qu'au niveau 1, le traitement est effectué sur un seul schéma qui correspond à la racine de la structure arborescente de notre stratégie de recherche hiérarchique (cf. § 5.6).

5.6.4.3 Conditions d'arrêts

Tout algorithme récursif nécessite une condition d'arrêt pour lui permettre de s'arrêter et ainsi fournir ses résultats. Nous utilisons deux conditions d'arrêt.

5.6.4.3.1 Condition d'arrêt 1

Soit $M_n^{q'}(q) = [CR_{n-1}^{q'}(q-1), CR_{n-1}^{q'}(q)]$ le schéma de mise en correspondance de rang q au niveau n , obtenu à partir du schéma de rang q' au niveau $n-1$. Rappelons que les couples de référence $CR_{n-1}^{q'}(q-1) = (g_{n-1}^{q'}(q-1), d_{n-1}^{q'}(q-1))$ et $CR_{n-1}^{q'}(q) = (g_{n-1}^{q'}(q), d_{n-1}^{q'}(q))$ sont obtenus à partir du schéma de rang q' au niveau $n-1$. Soient $G_n^{q'}(q)$ et $D_n^{q'}(q)$ respectivement les listes gauche et droite des points de contour à partir desquels le schéma $M_n^{q'}(q)$ est construit. La liste $G_n^{q'}(q)$ contient alors tous les points de contour de l'image-ligne épipolaire gauche situés entre les points $g_{n-1}^{q'}(q-1)$ et $g_{n-1}^{q'}(q)$. De même, la liste $D_n^{q'}(q)$ contient tous les points de contour de l'image-ligne épipolaire gauche situés entre les points $d_{n-1}^{q'}(q-1)$ et $d_{n-1}^{q'}(q)$.

Si l'une des deux listes $G_n^{q'}(q)$ et $D_n^{q'}(q)$ est vide, aucun appariement n'est possible. L'algorithme n'engendre plus de nouveaux schémas à partir du schéma $M_n^{q'}(q)$. C'est le cas par exemple sur la figure 5.4 pour les deux premiers schémas du niveau 2 de recherche. En se référant à la structure arborescente de notre stratégie hiérarchique, le nœud correspondant au schéma $M_n^{q'}(q)$ est dit de type terminal.

5.6.4.3.2 Condition d'arrêt 2

Il existe un niveau n_{max} de recherche à partir duquel nous sommes sûr que tous les points de contours des deux images-ligne épipolaires sont sélectionnés et traités par la procédure de mise en correspondance. La valeur n_{max} peut être fixée en fonction du critère de sélection (cf. § 5.6.4.1) et du seuil bas S_b utilisé lors de l'extraction des points de contour (cf. § 2.6.2.3).

Puisque les points de contour sont sélectionnés en fonction de l'amplitude de leur dérivée, il existe un niveau n_{max} pour lequel :

$$\frac{Max_g}{2n_{max}} < S_b. \quad (5.7)$$

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

$$\frac{Max_g}{2S_b} < n_{max}. \quad (5.8)$$

Des relations similaires, pour déterminer la valeur n_{max} , 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 $-S_b$.

Pratiquement, la valeur maximale de l'amplitude de la dérivée est 255, le seuil bas étant fixé 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 sélectionner et de traiter tous les points de contour des deux images-ligne épipolaires par la procédure de mise en correspondance. Le risque est que plus on monte dans les niveaux, plus la taille des schémas de mise en correspondance diminue et donc plus on perd l'apport des contraintes globales qui sont à la base de notre méthode de mise en correspondance. Pour éviter ce problème, on peut ajuster n_{max} à des valeurs inférieures à 13. Ainsi, à chaque fois que l'algorithme hiérarchique rencontre un schéma de mise en correspondance du niveau n_{max} de recherche, celui-ci sera traité par le module d'appariement élémentaire hiérarchique sans sélection (cf. § 5.6.4.4), c'est-à-dire en appliquant la procédure de mise en correspondance sur l'ensemble des points de contours relatifs au schéma. Si ce module fournit de nouvelles paires appariées, celles-ci généreront au niveau $n_{max} + 1$ de nouveaux schémas à partir desquels aucun traitement ne sera effectué. En d'autres termes, une fois que le traitement d'un schéma du niveau n_{max} est fini, l'algorithme passe au traitement du schéma suivant, s'il existe, au même niveau. Les nœuds correspondant aux schémas issus d'un schéma du niveau n_{max} sont alors des nœuds feuille. En se référant à la structure arborescente de notre stratégie hiérarchique, les nœuds correspondant aux schémas de niveau n_{max} sont dits de type post-terminal. Dans le cas où le module d'appariement élémentaire hiérarchique sans sélection ne fournit aucune paire à partir d'un schéma du niveau n_{max} , celui-ci correspondra alors à un nœud feuille.

5.6.4.4 Module d'appariement élémentaire hiérarchique

Soit $M_n^{q'}(q) = [CR_{n-1}^{q'}(q-1), CR_{n-1}^{q'}(q)]$ le schéma de mise en correspondance de rang q au niveau n , obtenu à partir du schéma de rang q' au niveau $n-1$. Rappelons que les couples de référence $CR_{n-1}^{q'}(q-1) = (g_{n-1}^{q'}(q-1), d_{n-1}^{q'}(q-1))$ et $CR_{n-1}^{q'}(q) = (g_{n-1}^{q'}(q), d_{n-1}^{q'}(q))$ sont obtenus à partir du schéma de rang q' au niveau $n-1$. Soient $G_n^{q'}(q)$ et $D_n^{q'}(q)$ respectivement les listes gauche et droite des points de contour à partir desquels le schéma $M_n^{q'}(q)$ est construit. La liste $G_n^{q'}(q)$ contient alors tous les points de contour de l'image-ligne épipolaire gauche situés entre les points $g_{n-1}^{q'}(q-1)$ et $g_{n-1}^{q'}(q)$. De même, la liste $D_n^{q'}(q)$ contient alors tous les points de contour de l'image-ligne épipolaire droite situés entre les points $d_{n-1}^{q'}(q-1)$ et $d_{n-1}^{q'}(q)$.

Selon le résultat du test d'arrêt appliqué au schéma $M_n^{q'}(q)$, le module d'appariement élémentaire hiérarchique peut être avec ou sans sélection. Si le schéma correspond à un nœud de type normal, c'est le module d'appariement élémentaire hiérarchique avec sélection qui est appliqué. Celui-ci consiste, dans une première étape, à déterminer à partir des listes $G_n^{q'}(q)$ et $D_n^{q'}(q)$ les listes $GS_n^{q'}(q)$ et $DS_n^{q'}(q)$ contenant les points de contour les plus significatifs en utilisant le critère de sélection (cf. § 5.6.4.1). La procédure de mise en correspondance est ensuite appliquée aux listes $GS_n^{q'}(q)$ et $DS_n^{q'}(q)$. Dans le cas où le schéma $M_n^{q'}(q)$ correspond à un nœud post-terminal, on fait appel au module d'appariement élémentaire hiérarchique sans sélection. Celui-ci consiste à appliquer directement la procédure de mise en correspondance aux listes $G_n^{q'}(q)$ et $D_n^{q'}(q)$. Si, à l'issue du module d'appariement élémentaire hiérarchique avec ou sans sélection, on obtient une liste non vide de couples de points de contour appariés, on passe à l'étape de génération des nouveaux schémas de mise en correspondance au niveau $n + 1$ de recherche. Pour cela, on ajoute à cette liste les couples de référence $CR_{n-1}^{q'}(q-1)$ et $CR_{n-1}^{q'}(q)$ qui ont été utilisés pour définir le schéma $M_n^{q'}(q)$. Ces deux couples sont nécessaires à la création des nouveaux schémas de mise en correspondance au niveau $n + 1$, notamment le premier et le dernier schémas. Comme la liste des couples de référence doit être classée selon l'ordre des points de contour dans les images-ligne épipolaires, les couples $CR_{n-1}^{q'}(q-1)$ et $CR_{n-1}^{q'}(q)$ sont insérés respectivement en tête et en queue de la liste des points de contour appariés.

A l'issue du traitement du schéma $M_n^{q'}(q)$, on obtient donc la liste $CR_n^q = \{CR_n^q(k), k = 0, \dots, K_n^q\}$ des $K_n^q + 1$ couples de référence. Celle-ci permet alors de générer au niveau $n + 1$ de recherche la liste $M_{n+1}^q = \{M_{n+1}^q(k), k = 0, \dots, K_n^q\}$ des $K_n^q + 1$ nouveaux schémas où $M_{n+1}^q(k) = [CR_n^q(k-1), CR_n^q(k)]$ est le schéma de mise en correspondance de rang k au niveau $n + 1$, obtenu à partir des $(k-1)^{\text{ème}}$ et $k^{\text{ème}}$ couples de référence issus du traitement du schéma de rang q au niveau n . Le processus continue ensuite au niveau $n + 1$ en commençant par le premier schéma de mise en correspondance de la liste M_{n+1}^q .

Dans le cas où le module d'appariement élémentaire hiérarchique avec sélection ne fournit aucune nouvelle paire appariée à partir d'un schéma, celui-ci correspond alors à un nœud feuille dans l'arbre de recherche.

Si, à partir d'un schéma, le module d'appariement élémentaire hiérarchique avec sélection ne fournit pas de nouveaux appariements, nous continuons la recherche dans le même schéma au niveau supérieur. En fait, lorsque le module d'appariement élémentaire hiérarchique avec sé-

lection ne fournit pas de nouveaux appariements, cela signifie soit que l'une des listes des points de contour sélectionnés est vide soit tout simplement que la procédure de mise en correspondance n'a pas réussi à trouver d'appariement. Avec l'examen du schéma au niveau supérieur, nous donnons donc une chance à la fois à la procédure de sélection (*cf.* § 5.6.4.1) afin de fournir des listes de point de contours non vides et à la procédure de mise en correspondance pour trouver des appariements.

5.7 Application du schéma hiérarchique de mise en correspondance

Dans cette section, nous exposons les résultats de mise en correspondance obtenus en utilisant le schéma hiérarchique décrit ci-dessus. Nous présentons également une comparaison quantitative de ces résultats avec ceux obtenus par la méthode de mise en correspondance élémentaire en considérant l'ensemble des points de contour des images-ligne gauche et droite dans un même niveau. Cette comparaison est basée sur l'évaluation des solutions en termes de satisfaction des contraintes globales de mise en correspondance, c'est-à-dire les contraintes d'unicité, d'ordre et de continuité de la disparité. Nous utilisons alors une fonction d'évaluation traduisant ces contraintes en termes de coût. Elle est construite de telle sorte que la meilleure solution est celle qui a le plus faible coût. En fait, cette fonction d'évaluation n'est autre que, à un terme près, la fonction d'évaluation utilisée dans l'algorithme génétique pour la recherche d'une solution de mise en correspondance entre les points de contours des images-ligne gauche et droite (*cf.* équation 4.2). Les résultats, les comparaisons et les discussions que nous allons entreprendre dans cette section sont issus du traitement des séquences stéréoscopiques "Piéton" et "Approche d'un véhicule" (*cf.* paragraphes 3.8.2 et 3.8.3). Nous rappelons que ces deux séquences sont acquises avec le stéréoscope linéaire développé dans le cadre de la détection d'obstacles à l'avant d'un véhicule. Chaque couple d'images-ligne stéréoscopiques est traité indépendamment des autres. A l'issue de chaque traitement, une liste des paires de points de contour appariés est constituée. Chaque paire permet, grâce aux formules de triangulation, de déterminer la distance et la position d'un point de la scène réelle (*cf.* § 2.4).

5.7.1 Séquence stéréoscopique "Piéton"

La figure 5.5 montre le résultat de mise en correspondance obtenu par application de la méthode hiérarchique sur la séquence "Piéton", dont les images brutes sont présentées au paragraphe 3.8.2.1 (cf. figure 3.18). Le résultat du traitement par la méthode de mise en correspondance élémentaire en considérant l'ensemble des points de contour des images-ligne gauche et droite est illustré sur la figure 5.6. Nous rappelons que ces images-résultats s'interprètent de la manière suivante : l'axe horizontal représente l'abscisse dans la scène. Un point se trouvant sur le bord gauche de chaque image est situé à 5 mètres à gauche de l'axe optique du stéréoscope linéaire. Un point se trouvant sur le bord droit est situé à 5 mètres à droite de ce même axe. L'axe vertical représente le temps, celui-ci défilant du haut vers le bas de chaque image. La distance par rapport au stéréoscope est représentée en couleur. Celle-ci varie du bleu lorsque le piéton est éloigné du stéréoscope au rouge lorsqu'il est proche de ce dernier.

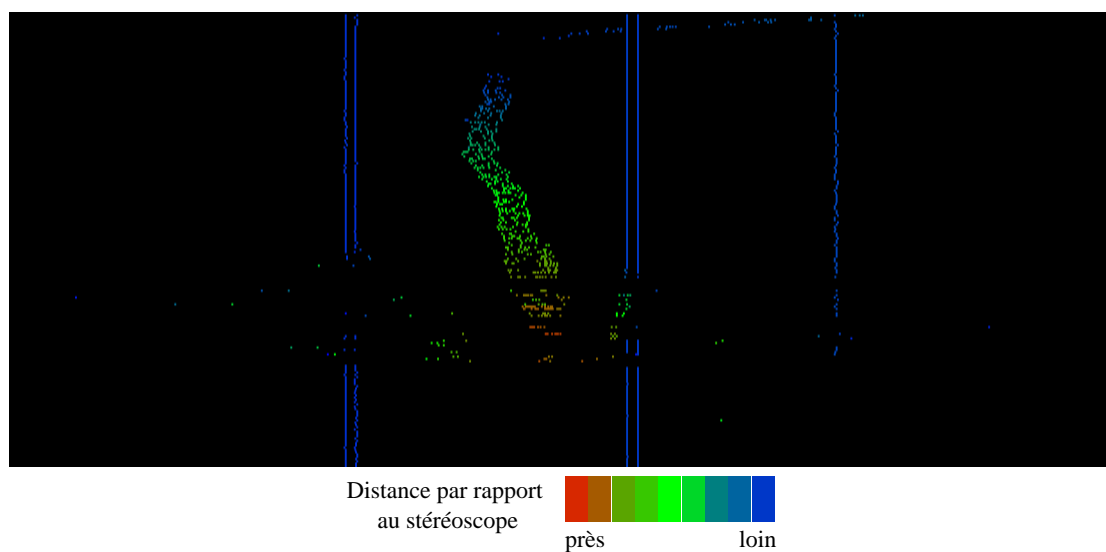


FIG. 5.5 : Résultat de mise en correspondance des couples d'images-ligne de la séquence "Piéton" par la méthode hiérarchique.

Pour les deux résultats, les points de contour des bandes blanches ainsi que ceux de la transition entre la chaussée et la zone d'ombre sont correctement appariés. Les points de contour correspondants au piéton sont également correctement appariés lorsqu'il n'y a pas de présence d'occlusions, c'est-à-dire lorsque le piéton ne cache pas l'une des bandes blanches à l'une des caméras. Sur la figure 5.6, on note la présence de mauvais appariements lorsque des occlusions se produisent, pratiquement entre les couples stéréoscopiques de rangs 100 et 150 de la séquence (cf. figure 4.11). Ces appariements erronés, qui sont peu nombreux, sont provoqués

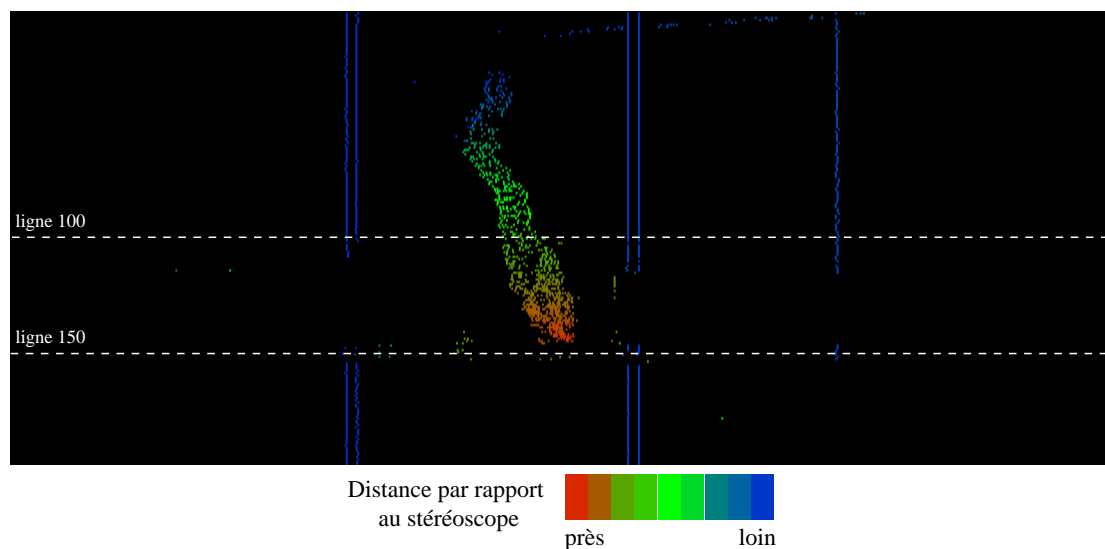


FIG. 5.6 : Résultat de mise en correspondance des couples d'images-ligne de la séquence "Piéton" par la méthode élémentaire.

par la mise en correspondance des points de contour d'une des bandes blanches avec ceux du piéton. Sur la figure 5.5, on remarque que, lorsque des occlusions se produisent, les appariements erronés sont plus nombreux. Afin de comparer de plus près les deux méthodes de mise en correspondance, nous procédons à une évaluation de leurs résultats en terme de satisfaction des contraintes d'unicité, d'ordre et de continuité de la disparité. Pour cela, nous nous servons d'une fonction d'évaluation qui représente ces contraintes en termes de coût. Elle est construite de telle sorte que la meilleure solution est celle qui correspond à un coût minimal. En fait, cette fonction d'évaluation n'est autre que, à un terme près, la fonction d'évaluation utilisée dans l'algorithme génétique de mise en correspondance des points de contour extraits des images-ligne gauche et droite (*cf.* équation 4.2). Étant donné que notre objectif est de comparer les résultats de mise en correspondance des deux méthodes par rapport aux trois contraintes stéréoscopiques, le dernier terme de l'équation 4.2 peut être supprimé. Nous rappelons que ce terme a été ajouté pour favoriser les chromosomes qui codent un nombre maximal d'appariements lors du déroulement de l'algorithme génétique. De ce fait, la fonction d'évaluation permettant de comparer les résultats de mise en correspondance des deux méthodes prend la forme :

$$\begin{aligned}
H = & K_{uni} \sum_{i=1}^{N_{min}} \left(1 - \sum_{j=1}^{N_{max}} U(i, C_j) \right)^2 + \\
& K_{ord} \sum_{i=1}^{N_{max}-1} \sum_{j=i+1}^{N_{max}} O(C_i, C_j) - \\
& K_{cont} \sum_{i=1}^{N_{max}-1} \sum_{j=i+1/C_i \neq C_j}^{N_{max}} S(C_i, C_j)
\end{aligned} \tag{5.9}$$

Nous rappelons que le premier terme correspond à la contrainte d'unicité où $U(i, C_j)$ représente une pénalité en cas de non respect de cette contrainte, c'est-à-dire lorsque le point de contour i d'une image-ligne est apparié avec plusieurs points de contour de l'autre image (cf. équation 5.9). Le second terme correspond à la contrainte d'ordre où $O(C_i, C_j)$ représente une pénalité en cas de non respect de cette contrainte entre les deux paires de points de contour (i, C_i) et (j, C_j) (cf. équation 5.9). Le troisième terme correspond à la contrainte de continuité de la disparité où $S(C_i, C_j)$ représente une mesure de compatibilité entre les deux paires de points de contour (i, C_i) et (j, C_j) (cf. équation 5.9). La figure 5.7 montre l'évaluation des résultats obtenus par les deux méthodes de mise en correspondance appliquées à la séquence "Piéton". L'axe horizontal représente le rang des couples d'images-ligne stéréoscopiques de la séquence et l'axe vertical représente l'évaluation du résultat de mise en correspondance.

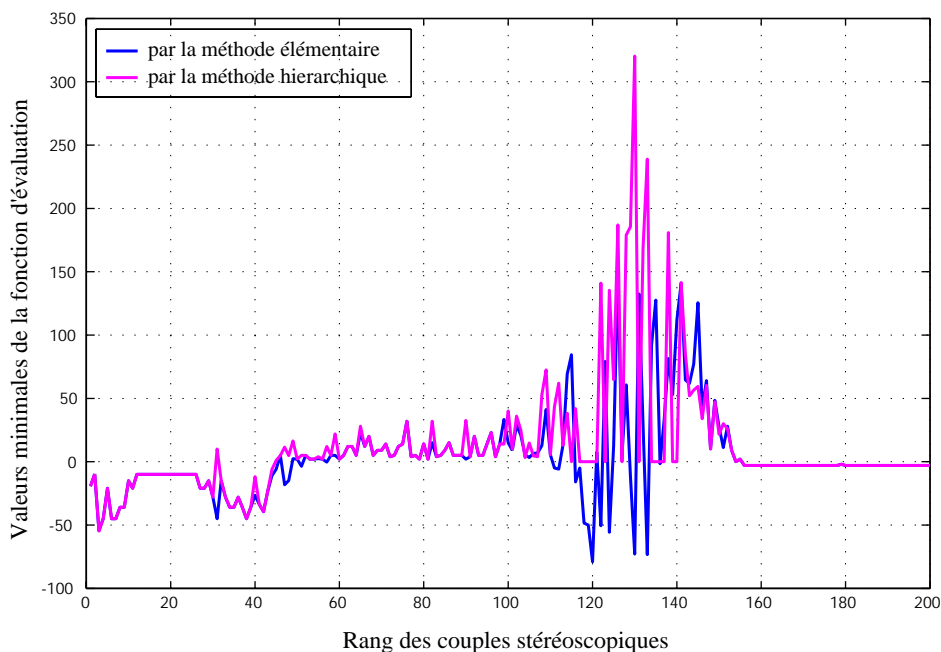


FIG. 5.7 : Évaluation des deux méthodes sur la séquence "Piéton".

La figure 5.7 met en évidence d'une manière quantitative les constatations que nous avons indiquées ci-dessus. Lorsque des occlusions se produisent, pratiquement entre les couples stéréoscopiques de rangs 100 et 150 de la séquence (cf. figure 4.11), on voit que la méthode hiérarchique se comporte mal et fournit des résultats de mise en correspondance moins bons que ceux obtenus par la méthode élémentaire. Sur les autres couples stéréoscopiques, c'est-à-dire quand il n'y a pas de problèmes d'occlusions, les deux méthodes se comportent pratiquement de la même façon et fournissent des résultats presque identiques.

Le mauvais comportement de la méthode hiérarchique peut s'expliquer par le fait que dans le schéma de recherche multi-niveaux, on perd un peu l'aspect global sur lequel est basé le processus de mise en correspondance à travers l'application des contraintes d'ordre et de continuité de la disparité. En effet, l'application de ces contraintes, qui sont des contraintes globales, ne prend plus en compte tous les points de contour des images-ligne. Le problème des faux appariements peut provenir également, surtout en cas d'occlusions, de la sélection des points de contour à chaque niveau de mise en correspondance. En effet, on peut se trouver dans des situations où un point de contour est sélectionné dans une image-ligne alors que son correspondant dans l'autre image-ligne ne l'est pas. Ainsi, lors de la mise en correspondance des points sélectionnés, des faux appariements vont se produire, mais également vont se propager puisque la mise en correspondances des autres points de contour en dépend. Passons maintenant à la comparaison des temps de traitement de chaque méthode. Avec la méthode élémentaire de mise en correspondance, la séquence "Piéton", composée de 200 couples stéréoscopiques, est traitée en 110 secondes environ, c'est-à-dire avec une cadence de traitement d'environ 2 paires par seconde. Avec les mêmes conditions paramétriques, cette séquence est traitée en 15 secondes environ lorsque la méthode hiérarchique est appliquée, c'est-à-dire avec une cadence de traitement d'environ 13 paires par seconde.

5.7.2 Séquence stéréoscopique "Approche d'un véhicule"

La figure 5.8 montre le résultat de mise en correspondance issu de l'application de la méthode hiérarchique sur la séquence "Approche d'un véhicule", dont les images brutes sont présentées au paragraphe 3.8.3.1. Le résultat fourni par la méthode élémentaire est illustré sur la figure 5.9. On obtient de bons résultats de mise en correspondance avec les deux méthodes. Cependant, pour les mêmes raisons que celles évoquées ci-dessus, la méthode hiérarchique

rencontre quelques difficultés de mise en correspondance à la fin de la séquence, c'est-à-dire lorsque des occlusions se produisent : dans l'image-ligne gauche, la bande blanche de gauche est cachée par le véhicule qui dépasse le véhicule instrumenté alors qu'une partie de cette bande blanche est visible dans l'image-ligne droite ; l'ombre projeté par le véhicule qui se rabat à l'avant du véhicule instrumenté cache la ligne en pointillés gauche dans l'image-ligne droite alors qu'elle est visible dans l'image-ligne gauche. Ces cas d'occlusions sont visibles entres les couples stéréoscopiques de rangs 140 et 180 de la séquence (cf. figure 3.35).

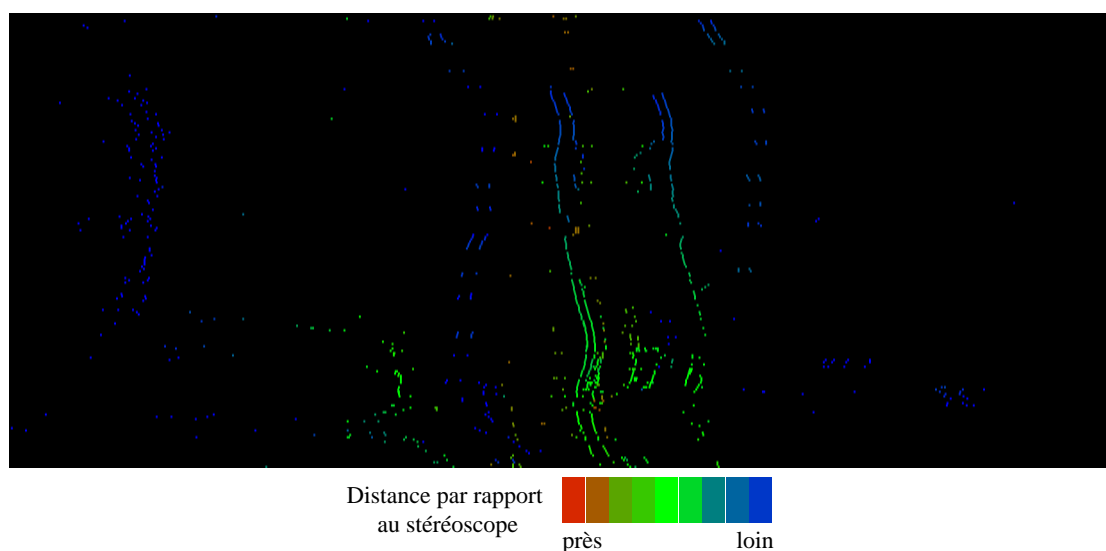


FIG. 5.8 : Résultat de mise en correspondance des couples d'images-ligne de la séquence "Approche d'un véhicule" par la méthode hiérarchique.

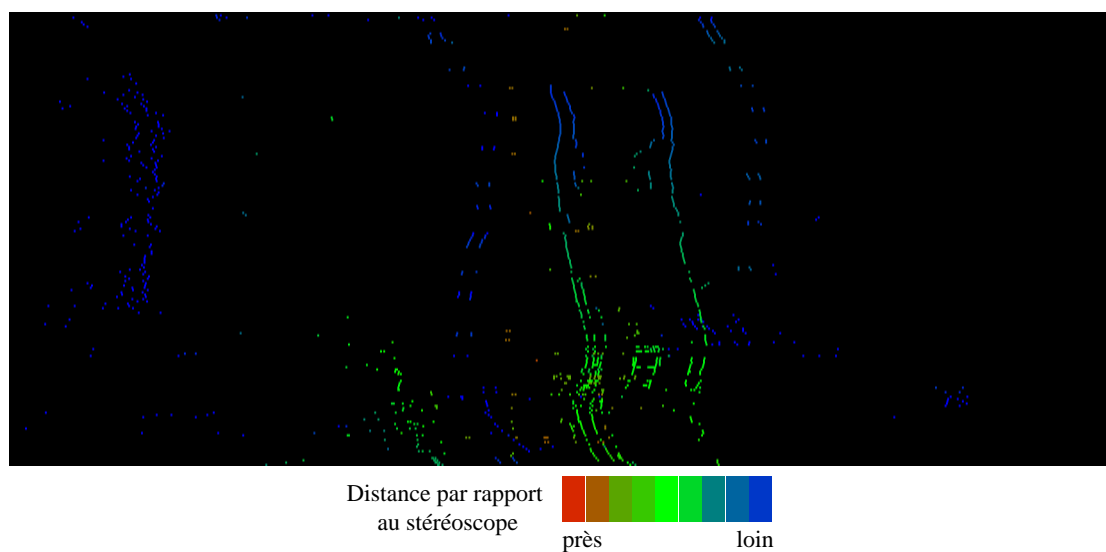


FIG. 5.9 : Résultat de mise en correspondance des couples d'images-ligne de la séquence "Approche d'un véhicule" par la méthode élémentaire.

La figure 5.10 montre l'évaluation quantitative des résultats obtenus par les deux méthodes de mise en correspondance appliquées à la séquence "Approche d'un véhicule". Comme nous l'avons mis en évidence précédemment, on peut constater que les deux méthodes se comportent à peu près de la même façon et fournissent des résultats presque identiques quand il n'y a pas d'occlusions, pratiquement du début de la séquence jusqu'au couple stéréoscopique de rang 140 et au delà du couple stéréoscopique de rang 180. Entre les couples stéréoscopiques de rangs 140 et 180 où des occlusions se produisent, on note le mauvais comportement de la méthode hiérarchique par rapport à celui de la méthode élémentaire.

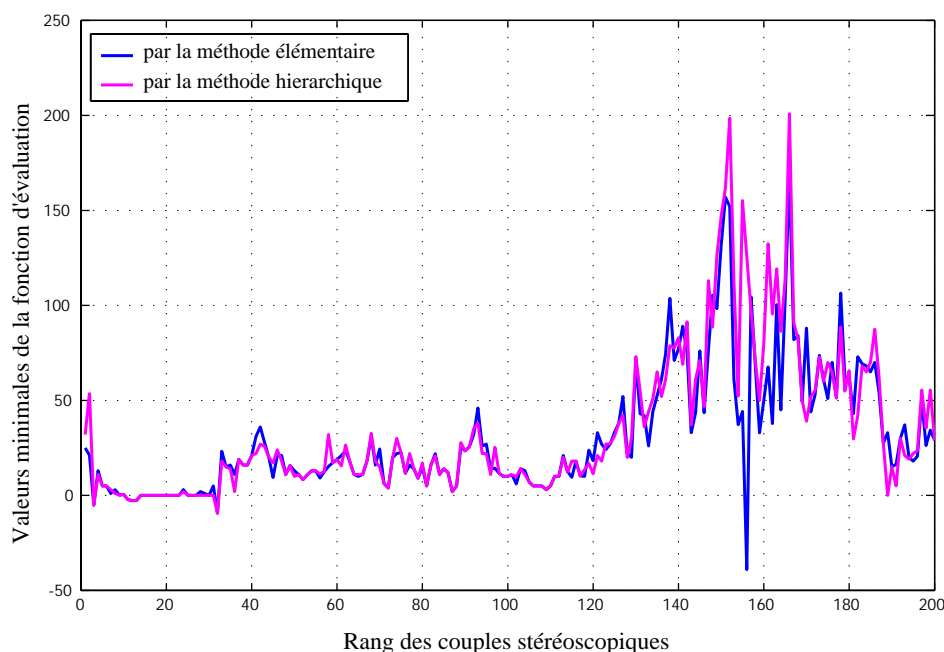


FIG. 5.10 : Évaluation des deux méthodes sur la séquence "Approche d'un véhicule".

Au niveau des temps de traitement, la méthode élémentaire traite la séquence "Approche d'un véhicule" en 70 secondes environ alors que la méthode hiérarchique n'a besoin que de 12 secondes environ pour effectuer ce traitement. Comme cette séquence est composée de 200 couples stéréoscopiques, la cadence de traitement avec la méthode élémentaire est d'environ 2.5 paires par seconde tandis qu'avec la méthode hiérarchique, environ 16 paires sont traitées par seconde.

5.7.3 Conclusion

L'analyse des résultats montre que les deux méthodes de mise en correspondance se comportent pratiquement de la même façon, sauf en cas d'occlusions où la méthode hiérarchique

rencontre quelques difficultés. Celles-ci sont dues au fait que, dans la méthode hiérarchique, les contraintes d'ordre et de continuité de la disparité, qui sont des contraintes globales, ne prennent pas en compte tous les points de contour des images. En effet, ces deux contraintes sont appliquées globalement, mais pour n'apparier à chaque fois que les points de contour sélectionnés dans chaque niveau de mise en correspondance. Le mauvais comportement de la méthode hiérarchique peut provenir également, surtout en cas d'occlusions, de la phase de sélection des points de contour à chaque niveau de mise en correspondance. En effet, lors de la mise en correspondance des points sélectionnés, des faux appariements vont se produire mais également vont se propager puisque la mise en correspondances des autres points de contour en dépend. En revanche, la recherche multi-niveaux permet de réduire la combinatoire du problème de mise en correspondance en le décomposant en plusieurs sous-problèmes de tailles moins importantes. La méthode hiérarchique permet ainsi un traitement plus rapide par rapport à la méthode élémentaire, ce qui permet d'envisager une exploitation effective du procédé.

5.8 Conclusion

L'inconvénient de la procédure de mise en correspondance élémentaire, présentée dans le chapitre précédent, est qu'elle nécessite un coût de calcul très important. Cette limitation est due principalement à la combinatoire élevée du problème de mise en correspondance. Afin d'envisager une exploitation effective de cette procédure, nous avons présenté dans ce chapitre une méthode de mise en correspondance hiérarchique permettant de décomposer le problème en plusieurs sous-problèmes de combinatoires moins élevées. Cette méthode consiste à effectuer une recherche des appariements par niveau. Pour un niveau de recherche donné, on commence par sélectionner les points de contour les plus significatifs de chaque images-ligne en utilisant l'information du gradient. Les points de contour sélectionnés sont ensuite appariés pour obtenir des couples de référence. Ces derniers vont générer des nouveaux plans de recherche dans lesquels la recherche se poursuit pour mettre en correspondance les points de contour moins significatifs dans le niveau de recherche suivant. La comparaison entre la méthode élémentaire et la méthode hiérarchique a montré que les deux méthodes fournissent pratiquement des résultats similaires sauf en cas d'occlusions où la recherche multi-niveaux est moins efficace. En revanche, la cadence de traitement de la méthode hiérarchique est plus intéressante que celle de la méthode élémentaire, ce qui permettrait d'envisager une exploitation effective de la procédure

de mise en correspondance pour des applications temps réel telles que la détection d'obstacles.

Chapitre 6

Conclusion générale

L'un des objectifs de la vision artificielle est de percevoir le relief d'un environnement à partir d'au moins deux images. Les recherches dans ce domaine ont permis de développer de nombreuses approches. La stéréovision constitue l'une des méthodes les plus répandues et les plus utilisées pour aborder ce problème. Il s'agit de retrouver la géométrie spatiale d'une scène à partir de plusieurs images de cette même scène, prises sous des angles différents.

Les différentes méthodes développées pour l'appariement d'images en stéréovision peuvent être regroupées en deux catégories : *méthodes denses* et *méthodes non-denses*. Dans cette thèse, nous nous sommes intéressés à la deuxième catégorie de méthodes en utilisant des primitives de type point de contour. Pour extraire ces points de contour on utilise le filtre de Deriche suivi d'une méthode de sélection proposée par Burie.

Afin d'apparier les points de contour extraits des images stéréoscopiques, nous avons présenté une approche globale basée sur des concepts de calcul génétique. Cette approche considère l'appariement des primitives extraites d'un couple stéréoscopique comme un problème d'optimisation où une fonction d'évaluation, intégrant des contraintes stéréoscopiques, doit être minimisée. Pour ce faire, nous avons élaboré un codage approprié permettant de représenter les solutions du problème sous forme de chromosomes qui, grâce à la prise en compte des contraintes locales de position et de signe de la dérivée, permettent de ne coder que des solutions compatibles vis-a-vis de celles-ci. Ensuite, nous avons défini une fonction d'évaluation permettant de juger de la qualité des chromosomes par rapport à l'objectif recherché, à savoir minimiser les erreurs commises en mettant en correspondance les primitives extraites de deux images stéréoscopiques. Nous avons ensuite décrit l'algorithme utilisé ainsi que les différents opérateurs génétiques nécessaires à son déroulement.

Pour valider notre approche, nous l'avons appliquée à des séquences d'images-ligne sté-

réoscopiques réelles acquises grâce à deux stéréoscopes développés dans le cadre de projets régionaux et européens dans lesquels le laboratoire I^3D est l'un des acteurs. Le premier stéréoscope, constitué de deux caméras linéaires, est utilisé pour la détection d'obstacles à l'avant d'un véhicule routier tandis que le deuxième, qui est constitué de deux caméras matricielles, est installé sur une rame du métro de Lille pour le même objectif.

Avec un simple codage binaire, les résultats de mise en correspondance obtenus sont encourageants. En effet, l'approche génétique développée fournit des résultats satisfaisants, notamment quand il s'agit d'images-ligne contenant un nombre limité de points de contour. Mais, quand ce nombre augmente, la taille des chromosomes générés avec le codage binaire devient importante, ce qui nécessite une augmentation de la taille de la population de chromosomes et du nombre de générations afin que l'algorithme génétique puisse converger vers des solutions acceptables. Cette augmentation des valeurs des deux paramètres génétiques se traduit par une augmentation considérable du temps de traitement nécessaire pour la phase de mise en correspondance.

Pour pallier cette limitation, nous avons proposé un autre codage. Les nouveaux chromosomes ainsi créés sont composés de gènes prenant des valeurs entières plutôt que binaires. Ce nouveau codage a permis, d'une part, de réduire la taille des chromosomes et, d'autre part, d'explorer plus efficacement l'espace de recherche de solutions grâce à l'élimination des gènes représentant des appariements ambigus. Ainsi, la probabilité de converger vers des solutions acceptables a été augmentée.

En dehors de ces deux avantages, le nouveau codage a permis de réduire d'un facteur 1.5 le temps de traitement par rapport au temps nécessaire dans le cas du codage binaire pour une séquence représentant un piéton traversant la chaussée. Cependant, ce temps reste assez important car l'évaluation de l'adaptation des chromosomes entiers s'effectue toujours à partir des chromosomes binaires correspondants. Nous avons alors modifié la fonction d'évaluation pour qu'elle permette de mesurer directement l'adaptation des chromosomes entiers sans utiliser leurs correspondants binaires. La nouvelle fonction d'évaluation a permis de réduire le temps de traitement d'un facteur 10 par rapport à celui qui est nécessaire en utilisant la fonction d'évaluation des chromosomes binaires. En d'autres termes, le codage entier et la nouvelle fonction d'évaluation ont permis, d'une part, d'explorer plus efficacement l'espace de solutions et, d'autre part, de réduire le temps de traitement d'environ un facteur 15 par rapport au temps

nécessaire dans le cas du codage binaire (0.6 seconde environ par couple d'images-ligne au lieu de 9 secondes).

L'approche génétique globale de mise en correspondance a été également appliquée pour appairer des primitives extraites des images-ligne d'une séquence stéréoscopique typique de scènes routières. Pour cette séquence, le temps de traitement a été évalué à 4 secondes environ par couple d'images-ligne dans le cas du codage binaire. Grâce au codage entier et à la nouvelle fonction d'évaluation, ce temps est passé à 0.45 secondes par couple d'images-ligne, ce qui signifie une réduction du temps de traitement d'un facteur proche de 8.5.

Ensuite, le codage entier et la nouvelle fonction d'évaluation ont été utilisés pour mettre en correspondance des primitives extraites de différents couples d'images stéréoscopiques de synthèse et réelles. Les résultats obtenus sur ces couples sont satisfaisants, mais le temps de traitement nécessaire pour l'appariement de certains couples d'images-ligne épipolaires reste trop important, notamment pour des applications nécessitant de réponses en "temps-réel" telle que la détection d'obstacles. En effet, lorsque le nombre de points de contour candidats à l'appariement augmente, la taille des chromosomes entiers augmente également. De ce fait, la convergence vers des solutions acceptables nécessite d'accroître la taille de la population de chromosomes et du nombre de générations, ce qui implique une augmentation du temps de calcul nécessaire pour la phase de mise en correspondance.

Afin d'envisager une exploitation effective de notre procédure génétique et palier la limitation due principalement à la combinatoire élevée du problème de mise en correspondance, nous avons présenté une méthode hiérarchique permettant de décomposer le problème en plusieurs sous-problèmes de combinatoires moins élevées. Cette méthode consiste à effectuer une recherche des appariements par niveau. Pour un niveau de recherche donné, on commence par sélectionner dans chaque image-ligne les points de contour les plus significatifs en utilisant l'information apportée par la valeur du gradient. Les points de contour sélectionnés sont ensuite appariés pour obtenir des couples de référence. Ces derniers gèrent de nouveaux plans de recherche dans lesquels il s'agit de mettre en correspondance les points de contour moins significatifs. La comparaison entre la méthode élémentaire et la méthode hiérarchique a montré que les deux méthodes fournissent pratiquement des résultats similaires, sauf en cas d'occlusions où la recherche multi-niveaux est moins efficace. En revanche, la cadence de traitement de la méthode hiérarchique est plus intéressante que celle de la méthode élémentaire, ce qui permet-

trait d'envisager une exploitation effective de la procédure de mise en correspondance pour des applications temps réel telles que la détection d'obstacles. Avec la méthode hiérarchique, les cadences de traitement sont d'environ 13 couples d'images-ligne par seconde pour la séquence "Piéton" et de 16 couples par seconde pour la séquence "Approche d'un véhicule".

Perspectives

Le travail décrit dans cette thèse devra être poursuivi dans deux directions principales. Comme nous l'avons montré au chapitre 3, une voie de recherche pour développer des AGs adaptatifs est très prometteuse. Il s'agit de trouver des critères et/ou des méthodes de réglage des paramètres de l'algorithme génétique en fonction de la complexité des images en termes de nombre de primitives. En effet, l'AG que nous avons proposé nécessite encore quelques améliorations, notamment au niveau du réglage des valeurs des paramètres génétiques qui régissent son fonctionnement. Dans ce cadre, nous pensons qu'un système de réglage dynamique plus approprié, notamment pour les valeurs des paramètres $Taille_{pop}$ et N_{gen} qui dépendraient fortement du nombre de points candidats à l'appariement, permettrait de converger plus rapidement vers des solutions satisfaisantes. D'autre part, nous n'avons pas eu l'occasion d'essayer d'autres opérateurs de croisement tels que, par exemple, le croisement en deux points ou le croisement uniforme. Une étude comparative des résultats de mise en correspondance obtenus en employant différents opérateurs de croisement serait particulièrement intéressante. Par ailleurs, l'implantation des procédures sur des architectures parallèles spécifiques permettrait d'améliorer encore plus les temps de traitement en utilisant plusieurs processeurs. L'un des modèles connus dans ce domaine est le modèle maître/esclaves. Selon ce modèle, le processeur maître assure la gestion des esclaves : chaque esclave reçoit son travail du maître, l'exécute, puis retourne son résultat au maître. Quand un processeur esclave devient disponible, le processeur maître lui envoie une autre tâche à exécuter. Notre approche, basée sur les concepts de calcul génétique, peut être aisément implantée sur ce type de modèle en attribuant, par exemple, les tâches de reproduction (croisement et mutation) et d'évaluation des chromosomes générés à plusieurs processeurs (esclaves) plutôt qu'à un seul.

Une deuxième voie intéressante concerne les primitives-image candidates à l'appariement. Comme nous l'avons montré précédemment, le problème de mise en correspondance est de nature combinatoire. Cette combinatoire devient élevée quand les primitives à apparier sont

nombreuses. Nous pensons qu'une analyse spatio-temporelle des séquences d'images stéréoscopiques permettrait de propager des hypothèses de mise en correspondance entre les couples d'images successifs. Ceci permettrait de réduire significativement l'espace de recherche de l'algorithme génétique.

Table des figures

2.1	Configuration générale d'un capteur stéréoscopique.	18
2.2	Configuration particulière du système de prise de vue.	19
2.3	Localisation tridimensionnelle.	21
2.4	Configuration du stéréoscope linéaire (les deux plans de visée sont confondus).	23
2.5	Triangulation géométrique dans le plan $(O, (X, Z))$	24
2.6	Cas particulier permettant de déterminer la coordonnée verticale d'un point vu par de caméras linéaires	25
2.7	Comparaison des primitives de type contour.	26
2.8	Extraction des extremums par intervalles de signe constant.	34
2.9	Les pixels b_g et b_d , projections de la mouche B , sont fortement corrélés, à la différence des pixels a_g et a_d , projections de A , qui reçoivent leur illumination de deux points physiques différents de la scène.	44
3.1	Croisement en position 2 de deux chromosomes parents de longueur 8.	51
3.2	Croisement uniforme.	52
3.3	Représentation schématique d'une mutation dans un chromosome binaire.	53
3.4	Schéma de mise en correspondance des points de contour d'un couple d'images-ligne épipolaires gauche et droite.	55
3.5	Tableau de mise en correspondance des listes des indices des points de contour G et D	56
3.6	Chromosome formé à partir du tableau de mise en correspondance.	56
3.7	Contrainte de position.	58
3.8	Contrainte de signe de la dérivée.	58
3.9	La contrainte d'ordre	59
3.10	Cas particulier où la contrainte d'ordre n'est pas respectée	59

3.11	Éléments parmi lesquels un chromosome est formé.	61
3.12	Exemple d'un chromosome compacte pour la mise en correspondance d'un couple d'images-ligne épipolaires avec $N_G = 6$ et $N_D = 5$	62
3.13	Codage de la contrainte d'unicité	63
3.14	Codage de la contrainte d'ordre	64
3.15	Graphe de la fonction non linéaire $S(X)$	65
3.16	Schéma de l'algorithme génétique.	67
3.17	Configuration du stéréoscope linéaire (Vue de côté).	72
3.18	Description de la scène "Piéton".	73
3.19	Déplacement d'un piéton vu par le stéréoscope linéaire.	73
3.20	Extraction des contours de la séquence "Piéton".	74
3.21	Résultat de la mise en correspondance de la séquence stéréoscopique "Piéton" grâce au codage binaire avec 50 chromosomes et 300 générations.	75
3.22	Nombre de points de contour extraits des images-ligne de la séquence "Piéton".	77
3.23	Évolution de la fonction d'évaluation pour le meilleur chromosome avec $Taille_{pop} = 50$ et $N_{gen} = 300$ pour le couple d'images-ligne épipolaires de rang 89 de la séquence "Piéton".	77
3.24	Évolution de la fonction d'évaluation pour le meilleur chromosome avec $Taille_{pop} = 50$ et $N_{gen} = 300$ pour le couple d'images-ligne épipolaires de rang 134 de la séquence "Piéton".	78
3.25	Résultat de la mise en correspondance de la séquence stéréoscopique "Piéton" pour différentes valeurs de $Taille_{pop}$ et de N_{gen}	79
3.26	Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 100$ et $N_{gen} = 300$ pour le couple de rang 89 de la séquence "Piéton". (Le changement d'échelle est effectué pour mettre en évidence l'évolution de la fonction d'évaluation au cours des dernières générations).	80
3.27	Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 50$ et $N_{gen} = 1000$ pour le couple de rang 89 de la séquence "Piéton". (Le changement d'échelle est effectué pour mettre en évidence l'évolution de la fonction d'évaluation au cours des dernières générations).	80

3.28	Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 100$ et $N_{gen} = 1000$ pour le couple de rang 89 de la séquence "Piéton". (Le changement d'échelle est effectué pour mettre en évidence l'évolution de la fonction d'évaluation au cours des dernières générations).	81
3.29	Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 100$ et $N_{gen} = 300$ pour le couple de rang 134 de la séquence "Piéton". (Le changement d'échelle est effectué pour mettre en évidence l'évolution de la fonction d'évaluation au cours des dernières générations).	81
3.30	Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 50$ et $N_{gen} = 1000$ pour le couple de rang 134 de la séquence "Piéton". (Le changement d'échelle est effectué pour mettre en évidence l'évolution de la fonction d'évaluation au cours des dernières générations).	82
3.31	Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 100$ et $N_{gen} = 1000$ pour le couple de rang 134 de la séquence "Piéton". (Le changement d'échelle est effectué pour mettre en évidence l'évolution de la fonction d'évaluation au cours des dernières générations).	82
3.32	Valeurs minimales de la fonction d'évaluation pour chacun des couples d'images-ligne de la séquence "Piéton" avec d'une part $Taille_{pop} = 50$ et $N_{gen} = 300$, et d'autre part $Taille_{pop} = 50$ et $N_{gen} = 1000$	83
3.33	Valeurs minimales de la fonction d'évaluation pour chacun des couples d'images-ligne de la séquence "Piéton" avec d'une part $Taille_{pop} = 100$ et $N_{gen} = 300$, et d'autre part $Taille_{pop} = 100$ et $N_{gen} = 1000$	83
3.34	Temps de traitement de la séquence "Piéton" pour différentes valeurs de $Taille_{pop}$ et de N_{gen}	84
3.35	Séquence stéréoscopique "Approche d'un véhicule".	85
3.36	Déplacements des différents véhicules au cours de la séquence. Les flèches indiquent les mouvements relatifs des véhicules par rapport au véhicule prototype.	85
3.37	Extraction des points de contours de la séquence "Approche d'un véhicule".	86

3.38	Résultat de la mise en correspondance de la séquence stéréoscopique "Approche d'un véhicule" obtenu grâce au codage binaire avec 100 chromosomes et 300 générations.	88
3.39	Nombre de points de contour extraits d'images-ligne de la séquence "Approche d'un véhicule".	89
3.40	Valeurs minimales atteintes de la fonction d'évaluation pour chacun des couples d'images-ligne de la séquence "Approche d'un véhicule" avec d'une part $Taille_{pop} = 50$ et $N_{gen} = 300$ et d'autre part $Taille_{pop} = 50$ et $N_{gen} = 1000$	90
3.41	Valeur minimales atteintes de la fonction d'évaluation pour chacun des couples d'images-ligne de la séquence "Approche d'un véhicule" avec d'une part $Taille_{pop} = 100$ et $N_{gen} = 300$ et d'autre part $Taille_{pop} = 100$ et $N_{gen} = 1000$	90
3.42	Temps de traitement de la séquence "Approche d'un véhicule" pour différentes valeurs de $Taille_{pop}$ et de N_{gen}	91
4.1	Chromosome entier constitué de N_{max} gènes. La valeur du gène i , $i \in T_{max} = \{1, 2, \dots, N_{max}\}$ est une valeur entière C_i , $C_i \in \{0 \cup T_{min}\}$ avec $T_{min} = \{1, \dots, N_{min}\}$	95
4.2	Exemple d'un chromosome généré grâce au codage entier où $i \in T_{max} = \{1, 2, 3, 4, 5, 6\}$ et $C_i \in \{0 \cup T_{min}\}$ avec $T_{min} = \{1, 2, 3, 4, 5\}$	95
4.3	Exemple illustrant : (a) un chromosome entier et le chromosome binaire équivalent, (b) les points de contour appariés parmi ceux extraits d'un couple de lignes épipolaires gauche et droite avec $N_G = 6$ et $N_D = 5$ et (c) le tableau binaire de mise en correspondance.	96
4.4	Exemple d'un schéma de mise en correspondance d'un couple d'images-ligne épipolaires avec $N_G = 15$ et $N_D = 17$	97
4.5	Exemple d'un chromosome entier codant une possibilité d'appariement d'un couple d'images-ligne épipolaires et le schéma de mise en correspondance équivalent.	98
4.6	Résultat de la mise en correspondance de la séquence "Piéton" avec chromosomes entiers évalués à partir de leur équivalent binaire.	101

4.7	Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 100$ et $N_{gen} = 300$ pour le couple d'images-ligne de rang 134 de la séquence "Piéton".	102
4.8	Résultat de la mise en correspondance de la séquence "Piéton" avec le codage entier et la nouvelle fonction d'évaluation.	105
4.9	Évolution de la fonction d'évaluation du meilleur chromosome avec $Taille_{pop} = 100$ et $N_{gen} = 300$ pour le couple d'images-ligne de rang 134 de la séquence "Piéton".	106
4.10	Résultat de la mise en correspondance de la séquence stéréoscopique "Approche d'un véhicule".	107
4.11	Séquence de déplacement d'un piéton.	108
4.12	Évolution de la fonction d'évaluation du meilleur chromosome pour quatre couples d'images-ligne extraits de la séquence "Piéton".	109
4.13	Évolution de la fonction d'évaluation du meilleur chromosome pour le couple d'images-ligne de rang 3 avec $N_G = 10$ et $N_D = 10$	110
4.14	Évolution de la fonction d'évaluation du meilleur chromosome pour le couple d'images-ligne de rang 88 avec $N_G = 10$ et $N_D = 10$	110
4.15	Évolution de la fonction d'évaluation du meilleur chromosome pour le couple d'images-ligne de rang 147 avec $N_G = 5$ et $N_D = 11$	111
4.16	Évolution de la fonction d'évaluation du meilleur chromosome pour le couple d'images-ligne de rang 150 avec $N_G = 5$ et $N_D = 11$	111
4.17	Couple d'images "Voie".	113
4.18	Extraction de contours du couple d'images "Voie".	114
4.19	Résultat de la mise en correspondance du couple d'images "Voie".	115
4.20	Couple d'images "Couloir".	116
4.21	Extraction de contours du couple d'images "Couloir".	116
4.22	Résultat de la mise en correspondance du couple d'images "Couloir"	117
4.23	Couple d'images "Village".	118
4.24	Extraction de contours du couple d'images "Village".	118
4.25	Résultat de la mise en correspondance du couple d'images "Village"	119
4.26	Stéréoscope dans la rame.	121

4.27	Couple d'images "Rails".	121
4.28	Extraction de contours du couple d'images "Rails".	121
4.29	Résultat de la mise en correspondance du couple d'images "Rails".	122
5.1	Schéma de mise en correspondance du problème initial.	127
5.2	Niveau 1 de recherche.	128
5.3	Niveau 2 de recherche.	129
5.4	Arbre de recherche de la stratégie hiérarchique.	133
5.5	Résultat de mise en correspondance des couples d'images-ligne de la sé- quence "Piéton" par la méthode hiérarchique.	143
5.6	Résultat de mise en correspondance des couples d'images-ligne de la sé- quence "Piéton" par la méthode élémentaire.	144
5.7	Évaluation des deux méthodes sur la séquence "Piéton".	145
5.8	Résultat de mise en correspondance des couples d'images-ligne de la sé- quence "Approche d'un véhicule" par la méthode hiérarchique.	147
5.9	Résultat de mise en correspondance des couples d'images-ligne de la sé- quence "Approche d'un véhicule" par la méthode élémentaire.	147
5.10	Évaluation des deux méthodes sur la séquence "Approche d'un véhicule". . .	148

Publications

1. Yassine Ruichek, Hazem Issa et Jack-Gérard Postaire. Genetic Approach for Obstacle Detection using Linear Stereo Vision. *IEEE Intelligent Vehicule Symposium*, Pages 261-266, Detroit, USA, Octobre 2000.
2. Hazem Issa, Yassine Ruichek et Jack-Gérard Postaire. A Genetic-Based Stereo Matching Technique. In *Proc. of the 2nd Int. Conf. on Philosophy and Computer Science - Processes of Evolution in Real and Visual Systems*, Krakow, Pologne, pages 101-105, Janvier 2002.
3. Hazem Issa, Yassine Ruichek et Jack-Gérard Postaire. A Genetic Optimization Scheme for Feature-Based Stereo Matching. In *Proc. of the 6th World Multiconference on Systemics, Cybernetics and Informatics*, volume 9, Orlando, USA, Juillet 2002.
4. Hazem Issa, Yassine Ruichek et Jack-Gérard Postaire. A Specific Encoding Scheme for Genetic Stereo Correspondence Searching : Application to Obstacle Detection. In *Journal of Systemics, Cybernetics and Informatics*, volume 1(1),
<http://www.iiisci.org/Journal/SCI/Contents.asp?var=&Previous=sci1>.
5. Hazem Issa, Yassine Ruichek et Jack-Gérard Postaire. Extracting Depth Information from Stereo Linear Images using a Genetic Approach. *IEEE 1st Int. Symposium on Intelligent Systems*, volume 1, pages 285-289, Varna, Bulgarie, Septembre 2002.
6. Hazem Issa, Yassine Ruichek et Jack-Gérard Postaire. A new Genetic Encoding for Edge Point Stereo Matching. In *Proc. of the 2nd Int. Conf. on signal, speech and Image Processing*, pages 169-174, Skiathos Island, Grec, Septembre 2002.
7. Hazem Issa, Yassine Ruichek et Jack-Gérard Postaire. Stereo Correspondence using a Genetic Scheme with a New Solution Encoding. *IEEE Int. Conf. on Systems Man and Cybernetics*, volume 6, IEEE catalog number : 02CH37349C, Hammamet, Tunisie, Octobre 2002.

8. Hazem Issa, Christophe Vieren, Yassine Ruichek et Louahdi Khoudour. Aide à l'exploitation des systèmes de transport guidés par stéréoscopie : une nouvelle approche par algorithme génétique. *Colloque Int. sur l'Innovation Technologique pour les Transports Terrestres*, volume 1, pages 305-312, Lille, France, Décembre 2003.

Bibliographie

- [AMM94] J. Arabas, Z. Michalewicz, et J. Mulawka. GAVaPS - a genetic algorithm with varying population size. In *Proc. of the 1st IEEE Conf. on Evolutionary Computation*, pages 73–78, 1994.
- [Aya89] N. Ayache. *Vision stéréoscopique et perception multisensorielle - Application à la robotique mobile*. Ed. InterEditions, Paris, 1989.
- [BBP95] J.-C. Burie, J.-L. Bruyelle, et J.-G. Postaire. Directing and localizing obstacles in front of a moving vehicle using linear stere vision. *Math. Comput. Modelling*, 22(4-7) :235–246, 1995.
- [Bäc92] T. Bäck. Self-adaptation in genetic algorithms. In *Proc. of the 1st European Conf. on Artificial Life*, pages 263–271, Cambridge, MA, 1992.
- [Bäc96] T. Bäck. *Evolutionary Algorithms in Theory and Practice*. Oxford University Press, Oxford, 1996.
- [Ben94] N. Benamrane. *Contribution à la vision stéréoscopique par mise en correspondance de régions*. Thèse de doctorat, Université de Valenciennes et du Hainaut Cambresis, France, 24 Janvier 1994.
- [Bru94] 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, France, Décembre 1994.
- [Bur95] 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, France, Septembre 1995.
- [Cab92] F. Cabestaing. *Détection de contours en mouvement dans une séquence d'images :*

- conception et réalisation d'un processeur cablé temps-réel*. Thèse de doctorat, Université des Sciences et Technologies de Lille, France, Janvier 1992.
- [Can86] J.-F Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6) :679–698, Novembre 1986.
- [CLRS99] P. Collet, E. Lutton, F. Raynal, et M. Schoenauer. Individual GP : an alternative viewpoint for the resolution of complex problems. In *Genetic and Evolutionary Computation Conf. GECCO99*, Orlando, Florida, USA, Juillet 1999.
- [CM99] Q. Chen et G. Medioni. Volumetric stereo matching method : Application to image-based modeling. In *Conf. on Computer Vision and Pattern Recognition*, volume 1, pages 29–34, Fort Collins, CO, USA, Juin 1999.
- [Cox94] I.J. Cox. A maximum likelihood N-camera stereo algorithm. In *Int. Conf. on Computer Vision and Pattern Recognition*, pages 733–739, 1994.
- [CPA95] J.M. Cruz, G. Pajares, et J. Aranda. A neural network model in stereovision matching. *Neural Networks*, 8(5) :805–813, 1995.
- [dBKS90] J.M.H. du Buf, M. Kardan, et M. Spann. Texture feature performance for image segmentation. *Pattern Recognition*, 23 :291–309, 1990.
- [Der87] R. Deriche. Using canny's criteria to derive a recursively implemented optimal edge detector. *Computer Vision*, 1(2) :167–187, Mai 1987.
- [Fa93] O.D. Faugeras et al. *Real time correlation-based stereo : algorithm, implementations and applications*. Rapport de recherche, Institut National des Sciences Appliquées de Lyon, INRIA Sophia-Antipolis, Août 1993.
- [Fua95] P. Fua. Reconstructing complex surfaces from multiple stereo views. In *Int. Conf. on Computer Vision*, pages 1078–1085, 1995.
- [GAF95] R.N. Greenwell, J.E. Angus, et M. Finck. Optimal mutation probability for genetic algorithms. *Math. Comput. Modelling*, 21(8) :1–11, 1995.
- [GL93] F. Glover et M. Laguna. Tabu search. In C. Reeves, editor, *Modern Heuristic Techniques for Combinatorial Problems*, pages 70–141, Oxford, England, 1993. Blackwell Scientific Publishing.
- [Gol89] D.E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, New York, 1989.

- [Gre86] J.J. Grefenstette. Optimization of control parameters for genetic algorithm. *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-16(1) :122–128, 1986.
- [Gri81] W.E.L. Grimson. A computer implementation of a theory of human stereo vision. In *Philosophical Trans. of the royal Society of London*, volume B 292, pages 217–253, May 1981.
- [Gri85] W.E.L. Grimson. Computational experiments with a feature based stereo algorithm. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 7(1) :17–34, January 1985.
- [GY02] M. Gong et Y.H. Yang. Genetic-based stereo algorithm and disparity map evaluation. *Computer Vision*, 47(1-3) :63–77, April 2002.
- [HL96] F. Herrera et M. Lozano. Adaptation of genetic algorithm parameters based on fuzzy logic controllers. *Genetic Algorithms and Soft Computing*, pages 95–125, 1996.
- [HLV97] F. Herrera, M. Lozano, et J.L. Verdegay. A learning process for fuzzy control rules using genetic algorithms. *Fuzzy Sets and Systems*, pages 143–158, 1997.
- [HM91] J. Hesser et R. Männer. Towards an optimal mutation probability in genetic algorithms. In *Parallel Problem Solving from Nature*, pages 23–32, 1991.
- [HM93] R. Horaud et O. Monga. *Vision par ordinateur - outils fondamentaux*. Ed. Hermès, Paris, 1993.
- [HM95] R. Horaud et O. Monga. *Vision par ordinateur*. Hermes, 2^{ème} édition, 1995.
- [HSC⁺89] W. Hoff, K.-W. Song, E.-Y. Chung, S.-J. Cho, et Y.-H. Ha. Surface from stereo : Integrating feature matching, disparity estimation, and contour detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(2) :121–136, February 1989.
- [HYFS96] C.C. Hsu, S.I. Yamada, H. Fujikawa, et K. Shida. A fuzzy self-tuning parallel genetic algorithm for optimization. *Computers ind. Enging.*, 30 :883–893, 1996.
- [Jon75] K.A. De Jong. An analysis of the behavior of a class of genetic adaptive systems. Ph.D. Dissertation 76-9381, University of Michigan, 1975.
- [KBL93] A. Khotanzad, A. Bokil, et Y.W. Lee. Stereopsis by constraint learning feed-

- forward neural networks. *IEEE Trans. on Neural Networks*, 4(9) :332–342, Mars 1993.
- [KHK90] E. Krotkov, K. Henriksen, et R. Kories. Stereo ranging with verging cameras. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(12) :1200–1205, Décembre 1990.
- [Kir84] S. Kirkpatrick. Optimization by simulated annealing : quantitative studies. *Journal of Statistical Physics*, 34(5-6) :975–986, 1984.
- [KO94] T. Kanade et M. Okutomi. Stereo matching algorithm with an adaptive window : Theory and experiment. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(9) :920–932, Septembre 1994.
- [KWT87a] M. Kass, A. Witkin, et D. Terzopoulos. Snakes : Active contour models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 3 :259–267, 1987.
- [KWT87b] M. Kass, A. Witkin, et D. Terzopoulos. Snakes : Active contour models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 3 :259–268, 1987.
- [Las96] P. Lasserre. *Vision pour la robotique mobile en environnement naturel*. PhD thesis, Université Paul Sabatier, Toulouse, France, Septembre 1996.
- [LGLB01] J. Louchet, M. Guyon, M.-J. Lesot, et A. Boumaza. L’algorithme des mouches dynamiques : Guider un robot par évolution artificielle en temps réel. *Apprentissage et évolution*, pages 115–130, 2001.
- [LL90] H.-J Lee et W.-L Lei. Region matching and depth finding for 3D objects in stereo aerial photographs. *Pattern Recognition*, 23(1) :81–94, 1990.
- [LL94a] S.-H. Lee et J.-J. Leou. A dynamic programming approach to line segment matching in stereo vision. *Pattern Recognition*, 27(8) :961–986, Août 1994.
- [LL94b] J.J. Leou et S.H. Lee. A dynamic programming approach to line segment matching in stereo vision. *Pattern Recognition Letters*, 27(8) :961–986, 1994.
- [Lou00] J. Louchet. Stereo analysis using individual evolution strategy. In *Int. Conf. on Pattern Recognition, ICPR2000*, pages 170–176, September 2000.
- [LOY73] M.D. Levine, D.A. O’Handley, et G.M. Yagi. Computer determination of depth maps. *Computer Graphics and Image Processing*, 2(4) :131–150, 1973.

- [LSH94] J.J. Lee, J.C. Shim, et Y.H. Ha. Stereo correspondence using the hopfield neural network of a new energy function. *Pattern Recognition*, 27(11) :1513–1522, 1994.
- [MD89] O. Monga et R. Deriche. 3D edge detection using recursive filtering : Application to scanner images. In *Int. Conf. on Vision and Pattern Recognition*, pages 28–35, 4-8 Juin 1989.
- [Mic94] Z. Michalewicz. *Genetic Algorithms + Data Structures = Evolution Programs*. Springer-Verlag, seconde édition, Berlin, 1994.
- [Mor77] H.-P. Moravec. Toward automatic visual obstacle avoidance. In *Proc. of the 5th Int. Joint Conf. on Artificial Intelligence*, page 584, Cambridge, MA, USA, Août 1977.
- [MP76] D. Marr et T. Poggio. Cooperative computation of stereo disparity. *Science*, 194 :283–287, 1976.
- [MTM93] F. Martinez, B. Thiesse, et P. Marthon. Reconstruction de scènes 3D en vision binoculaire animée. In *4^{ème} Journées ORASIS, GDR-PRC Communication Homme-Machine*, pages 13–14, Mulhouse, 11-14 Octobre 1993.
- [NC92] N.M. Nasrabadi et C.Y. Choo. Hopfield network for stereo vision correspondence. *IEEE Trans. on Neural Networks*, 3(1) :5–13, Janvier 1992.
- [OBMD94] V. Oksenhendler, A. Bensrhair, P. Miché, et R. Debrie. Un nouveau processus auto-adaptatif de stéréovision passive utilisant une nouvelle primitive : La déclivité. In *Colloque AGI'94*, pages 333–336, Futuroscope, Poitiers, 2-3 Juin 1994.
- [OK85] Y. Otha et T. Kanade. Stereo by intra- and inter-scanline search using dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7 :139–154, Mars 1985.
- [PM97] F. Pla et J.A. Marchant. Matching feature points in image sequences through a region-based method. *Computer Vision Image Understanding*, 66(3) :271–285, 1997.
- [RP99] Y. Ruichek et J.-G. Postaire. A new neural real-time implementation for obstacle detection using linear vision. *Real-Time Imaging Journal*, 5 :141–153, 1999.
- [Rui97] 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, France, 07 Février 1997.
- [SCED89] J.D. Schaffer, R.A. Caruna, L.J. Eshelman, et R. Das. A study of control parameters affecting online performance of genetic algorithms for function optimization. In *Proc. of the 3th Int. Conf. on Genetic Algorithms*, pages 51–60, 1989.
- [SH88] T. Skordas et R. Horaud. Mise en correspondance structurelle pour la vision stéréoscopique. In *Technique et Science Informatiques*, volume 7(6), pages 591–608, Lyon, France, 1988.
- [Sle95] M. Slesis. *Suivi et location 3D d'objets en mouvement à l'aide de modèles de contours actifs*. Thèse de doctorat, Université des Sciences et Technologies de Lille, France, 1995.
- [SM95] H. Saito et M. Mori. Application of genetic algorithms to stereo matching of images. *Pattern Recognition Letters*, 16(8) :815–821, August 1995.
- [SO94] K. Satoh et Y. Ohta. Passive depth acquisition for 3d image displays. *IEICE Trans. on Information and Systems*, E77-D(9) :949–957, September 1994.
- [SO95] K. Satoh et Y. Ohta. Occlusion detectable stereo using a camera matrix. In *Asian Conf. on Computer Vision, Int. Association for Pattern Recognition*, volume 1, pages 331–335, Singapore, 5-8 December 1995.
- [SP94] M. Srinivas et L. M. Patnaik. Adaptive probabilities of crossover and mutation in genetic algorithms. *IEEE Trans. on Systems, Man, and Cybernetics*, 24 :656–667, 1994.
- [Sto00] V. Stoclin. *Développement de critères pour l'évaluation de filtres discrets de détection de contours dans les images numériques*. Thèse de doctorat, Université des Sciences et Technologies de Lille, Décembre 2000.
- [VG91] L. Vinet et A. Gagalowicz. Mise en correspondance de régions utilisant la cohérence de leurs voisinages. In *8^{ème} Congrès de Reconnaissance des Formes et Intelligence Artificielle*, volume 2, pages 821–826, 25-29 Novembre 1991.
- [Vie88] C. Vieren. *Segmentation de scènes dynamiques en temps réel. Application au traitement de séquences d'images pour la surveillance de carrefours routiers*. Thèse de doctorat, Université des Sciences et Technologies de Lille, Avril 1988.

- [Wan96] 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, Décembre 1996.
- [WL95] M.S. Wu et J.J. Leou. A bipartite matching approach to feature correspondence in stereo vision. *Pattern Recognition Letters*, 16 :23–31, 1995.
- [Wu98] W. Wu. *Synthèse d'un contrôleur flou par Algorithme Génétique : Application au réglage dynamique des paramètres d'un système*. Thèse de doctorat, Université des Sciences et Technologies de Lille, Décembre 1998.
- [WWSJ96] P.Y. Wang, G.S. Wang, Y.H. Song, et A.T. Johns. Fuzzy logic controlled genetic algorithms. In *Proc. of the 5th IEEE Int. Conf. on Fuzzy Systems*, pages 972–979, 1996.
- [Xie95] M. Xie. Cooperative strategy for matching multilevel edge primitives. *Image Vision Computer*, 13(2) :89–99, 1995.
- [XV94] H.Y. Xu et G. Vukovich. Fuzzy evolutionary algorithms and automatic robot trajectory generation. In *Proc. of the 1st IEEE Conf. on Evolutionary Computation*, pages 595–600, 1994.
- [YKK86] M. Yachida, Y. Kitamura, et M. Kimachi. Trinocular vision : new approach for correspondence problem. In *Int. Conf. on Pattern Recognition*, pages 1041–1044, Paris, Octobre 1986.
- [Zio91] D. Ziou. *La détection de contours dans des images à niveaux de gris : mise en oeuvre et sélection de détecteurs*. Thèse de doctorat, Institut National Polytechnique de Lorraine, Octobre 1991.
- [ZK00] L.C. Zitnick et T. Kanade. A cooperative algorithm for stereo matching and occlusion detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(7) :675–684, 2000.

RÉSUMÉ en français

La stéréovision est une méthode très utilisée pour la perception du relief à partir d'images prises sous des angles différents. Le problème clé de cette méthode est l'appariement des primitives extraites des images stéréoscopiques. Dans cette thèse, nous présentons une approche globale pour l'appariement stéréoscopique de points de contour. Dans cette approche l'appariement stéréoscopique est considéré comme un problème d'optimisation dont la résolution est basée sur les concepts de calcul génétique. Grâce à un codage spécifique, une fonction d'évaluation des solutions est construite à partir des contraintes stéréoscopiques. Un algorithme génétique est ensuite mis en œuvre pour explorer l'espace des solutions. Afin de réduire la complexité du problème, nous proposons une stratégie hiérarchique pour la recherche des appariements. L'évaluation de nos algorithmes, appliqués sur de nombreuses images stéréoscopiques, montre l'intérêt de notre approche pour la reconstruction 3D d'une scène.

MOTS-CLÉS : Algorithmes génétiques ; Codages génétiques ; Extraction de primitives ; Mise en correspondance ; Recherche hiérarchique ; Reconstruction 3D ; Stéréovision.

RÉSUMÉ en anglais

Stereo vision is a well-known technique for obtaining depth information from two or more video images of a scene. The heavily investigated problem in this approach is how to find the corresponding features extracted from the stereo images. In this thesis, we present a new genetic based method for matching edge points extracted from stereo images. In our approach the matching problem is formulated as an optimisation task where an objective function, representing the constraints on the solution, is to be minimized thanks to a genetic algorithm with a specific encoding scheme. To reduce the complexity of the matching problem, the proposed genetic algorithm is applied according a hierarchical searching strategy. To demonstrate its effectiveness, the proposed genetic stereo matching algorithms are applied using several stereoscopic images.

KEYWORDS : Depth computation ; Feature extraction ; Genetic encoding schemes ; Genetic algorithms ; Hierarchical searching ; Stereo matching ; Stereo vision.